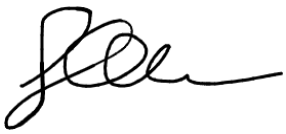
Exploring the role of diversity in bacteria-phage ecology and evolution

Submitted by

Jack Common

to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Biological Sciences
in July 2020

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: 

……..

Abstract

Host diversity is one of the key factors that determine why pathogens can emerge and cause epidemics, the scale of an epidemic, and host-pathogen coevolution. Although the effects of host diversity on pathogen spread are well-described, we currently lack a detailed understanding of the ecological and evolutionary mechanisms that underpin them. The interactions between phage and bacteria with a CRISPR-Cas immune system are a tractable model system to study how host density and genetic diversity impacts phage epidemiology, evolution and coevolution between the host and the phage. In this thesis, I first summarize existing literature on the role of host genetic diversity on host-parasite interactions. I then present experiments that explore how population bottlenecks, which impact both host density and genetic diversity, affect the epidemiology and evolution of phage. I find that bottlenecks result in more rapid phage extinction, which is driven primarily by the changes in host density. Next, I explore how manipulations of host genetic diversity alter the epidemiology and evolution of a phage that can infect a single host genotype in the population. I find that greater diversity results in protection of the susceptible host due to reduced contact rates with phage. I also find evidence for interaction between host diversity, phage population size, and the likelihood of phage evolution. Finally, I examine the coevolutionary interaction between phage and bacteria with CRISPR immune systems. I find that bacterial populations initially evolve diversity in their CRISPR-Cas immune systems that is sufficiently low for an escalating arms-race dynamic to ensue, that lasts until diversity in the CRISPR-Cas resistance alleles becomes too high, causing phage extinction.

Table of Contents

List of figures and tables

Figures

Tables

Acknowledgements

There are many people who have provided love, friendship, and guidance during this PhD. I thank everyone from the Westra lab for their help with lab issues and chats over coffee or a pint – Stineke, Sean, Anne, Clare, Devi, Bridget, John, and Mariann. I also thank the excellent technical staff, Daniela and Jenny, for ensuring things rang smoothly.

I thank my thesis examiners, Michael Brockhurst and Chris Lowe, for the helpful comments and invigorating discussion during the viva.

I want to thank my fellow students Ellinor, Ellie, Alice and Jenny for making me laugh and listening to me moan for four years. I particularly thank David and Dan for their contributions to chapters 3 and 4, respectively.

I am tremendously grateful to Edze Westra for having my back and believing in me during this PhD, as well as his significant contribution to the research that forms this thesis.

I thank my parents, Lisa and Chris, who were a safe harbour that steadied the ship when it threatened to capsize. I also thank Cath and Phil Sleigh, who provided a second home of warmth and light.

Finally, to Vita.

Author’s declaration

Chapters 2, 3 and 4 are the result of collaborative efforts, although the majority of the work is my own. Details of author contributions are outlined in the relevant section of chapter 1. Chapters 1 and 5 are entirely my own work, with critical revisions from Edze Westra.

Abbreviations

Unless otherwise indicated, chemical names follow IUPAC nomenclature.

BIM	**Bacteriophage-insensitive mutant**
Cfu	Colony-forming unit
Pfu	Plaque-forming unit
Phage	Bacteriophage
PWD	Pairwise difference
SM	*of bacterial immunity* surface modification *of a bacterial clone* surface mutant

Chapter 1 – Introduction

Literature review

Historical perspective

There are many examples of how the emergence and spread of infectious diseases have impacted and shaped histories, both large and small (Diamond, 2002). Mythological plagues, such as the pestilence and boils visited upon the people of Egypt by the god of the Old Testament, have long formed part of human ideas about divinity and the supernatural. The Plague of Athens in 430BCE, generally attributed to a typhoid fever (Papagrigorakis *et al.*, 2008; Littman, 2009), not only killed a significant proportion of Athenians, but changed the course of a war (Thucydides *et al.*, 1972) and inspired numerous works of Western art and literature. The Plague of Justinian (541-549 CE), caused by a strain of *Yersinia pestis* that swept through Europe and the Middle East (Harbeck *et al.*, 2013), permanently altered the balance of power in the Christian church and changed the course of European history (Rosen, 2010). The most well-known pandemic of *Y. pestis*, the Black Death, decimated the populations of Eurasia and Africa between 1347 and 1351, killing around a third of Europeans (Gould, 1966). The consequences of the Black Death were vast: it was a key trigger for the Hundred Years War between England and France (Turchin, 2018); it deepened antisemitism across Europe (Moore, 2008; Nirenberg, 2015); labour and land reforms in response to the vast loss of life realigned the feudal system, which in England led to the Peasant's Revolt (Dyer, 2002); and architecture and art moved toward greater realism and emotive

expression (Honour & Fleming, 2005). The use of a single-cultivar clone of potato that was susceptible to a strain of late blight fungus (*Phytomophthora infestans*) triggered the Irish potato famine (Machida-Hirano, 2015).

In the last century, three examples of infectious diseases stand out. The huge mortality of the 1918 flu pandemic, caused by a strain of H1N1 influenza A virus, was mainly due to the economic and political upheavals immediately following the First World War (1914-18), rather than the virus being unusually deadly (Brundage & Shanks, 2007; Morens & Fauci, 2007). The 1918 pandemic also greatly influenced the preparedness of public health systems toward another similarly infectious global flu outbreak. The ongoing HIV/AIDS pandemic, which as of 2018 killed around 32 million people (World Health Organisation, 2020), originated as SIV in non-human primates and adapted to infect humans in what is now the Democratic Republic of Congo sometime in the early 20th Century (Sharp & Hahn, 2011). The disproportionate impact of HIV/AIDS on the gay community and people in sub-Saharan African countries demonstrates how politics, colonialism, and prejudice can align to heighten the burden of infectious diseases on marginalised communities. Finally, the impacts of SARS-CoV-2, which likely emerged in a bat reservoir before crossing over to humans (Wu *et al.*, 2020), are still developing. This most recent pandemic has highlighted deep social inequalities, for example with severe cases and mortality disproportionately falling on black and brown communities in the UK as a result of systemic racism (Booth, 2020).

Observational studies

There is a breadth of evidence that crop diversity can prevent and/or mitigate the impact of pathogens, and that this can maintain or increase yields (Smithson & Lenne, 1996; Mundt, 2002; Borg *et al.*, 2018; Reiss & Drinkwater, 2018). Many subsistence peasant agriculturalists and indigenous peoples have long recognised and advocated the benefits of crop diversity (Mann, 2005; Chohan, 2017; Laughton, 2017). The benefits of cultivar mixtures were first experimentally studied in small grain crops prior to World War Two (Frankel, 1939), and later, work by Elton (1958) formalised the idea of a "monoculture effect" on disease susceptibility and transmission in crop species. The first systematic review of crop diversity showed some disease reduction benefits (Smithson & Lenne, 1996), and a later review by Mundt (2002) showed that cultivar mixes can effectively limit disease in small grain species. More recently, meta-analyses by Kiær *et al.* (2009) and Reiss and Drinkwater (2018) found that increased cultivar diversity based on disease traits led to significant yield increases. Reiss and Drinkwater (2018) note that the benefit of functional diversity in terms of disease was most important when other physical stressors are present. In other words, more diverse crop assemblages generate higher yields in the context of other abiotic perturbations compared to less diverse assemblages. This highlights the synergy between biotic and abiotic pressures, and is also particularly relevant given the expectation that both disease and physical pressure on crops is expected to increase with climate breakdown (Tripathi *et al.*, 2016). As changing climatic conditions continue to threaten the food supply, increasing crop diversity will become even more important.

Like crops, loss of genetic diversity in animal populations may render declining populations more vulnerable to disease. Numerous studies of wild host-pathogen systems have shown that host genetic diversity reduces pathogen prevalence and/or disease burden (O'Brien *et al.*, 1985; Thorne & Williams, 1988; Meagher, 1999; Acevedo-Whitehouse *et al.*, 2003; Pearman & Garner, 2005; Whiteman *et al.*, 2007; Ashby & King, 2015; van Houte *et al.*, 2016b; Eastwood *et al.*, 2017; López-Uribe *et al.*, 2017; Morley *et al.*, 2017). A recent meta-analysis provided broad support for a negative effect of host population genetic diversity on parasite success across the tree of life (Ekroth *et al.*, 2019). In this study, the authors noted that work in the field versus laboratory influenced the size of the effect, and that the success of microparasites (i.e. viruses, prokaryotes, and fungi) tended to be more affected by host diversity. This latter effect may be explained by physiological differences in how macro- and microparasites tend to interact with their hosts. Where macroparasites more often encounter physical and/or innate immune defences, such as skin thickness or an inflammatory response, microparasites often have more complex cellular or molecular interactions with the host immune system. As such, microparasites must evade or overcome host mechanisms with very high dimensionality and evolutionary potential, and so they may be more sensitive to diversity in such mechanisms.

The caveats to the diversity-disease effect highlighted by Ekroth *et al.* (2019) suggest that, while it is a useful heuristic that generally holds, specific host-pathogen systems in different contexts may show different responses. For example, in populations of the buff-tailed bumblebee (*Bombus terrestris*), the microsporidian parasite *Nosema bombi* is more successful in genetically homogeneous populations

compared to heterogeneous populations of *B. terrestris* (Baer & Schmid-Hempel, 1999; Baer & Schmid-Hempel, 2001, 2003), but the trypanosomelid *Crithidia bombi* is equally successful in both diverse and non-diverse hosts populations (Baer & Schmid-Hempel, 1999). In experimental populations of the ant *Formica selysi* infected with a single pathogen species, host genetic diversity does not always reduce pathogen success (Reber *et al.*, 2008; Schmidt *et al.*, 2011). Finally, because parasite success can be measured in different ways, the sign of the effect of host diversity - or indeed if an effect is detected at all – can vary between and within studies. Desai and Currie (2015) illustrate this in the honeybee *Apis mellifera* populations. Genetic diversity negatively affected pathogen success when infection prevalence or parasite load was calculated, but not when host survival was considered.

While the literature reviewed above clearly shows that diversity broadly benefits hosts, there is comparatively less clarity regarding *why* diversity is beneficial. In considering the ecological and evolutionary mechanisms that can drive the diversity-disease effect, the key questions to pose would be why emergence of infectious diseases occurs, and what determines their spread once they have emerged?

Theory

Early deterministic epidemiological models of pathogen spread assume that whether or not a pathogen spreads in a host population is due to the value of R_0, which is the expected number of secondary infections per primary case in a fully susceptible population (Anderson & May, 1992; Diekmann & Heesterbeek, 2000). In this

simplified scenario, an epidemic occurs if $R_0 > 1$, and if $R_0 < 1$, it decelerates or ends. Although the underlying assumptions of these kinds of deterministic models are limited, they have nonetheless been extremely useful in developing public health interventions and policies, such as vaccination programmes.

Simple probabilistic models which take into account host demographic stochasticity such as transmission, mortality and recovery rates, as well as the number of initially infected individuals show that pathogen extinction still occurs when $R_0 < 1$, even with multiple introduction events into the host population (Gandon *et al.*, 2013). When $R_0 > 1$ in such models, there is a positive non-linear relationship between R_0 and the number of introduced infected individuals. Conceivably, a pathogen with a sufficiently high R_0 can emerge in an uninfected host population from a single infected individual (May *et al.*, 2001).

Pathogens evolve, but epidemiological models like those discussed above tend to ignore pathogen evolution. Antia *et al.* (2003) illustrate how pathogens with an initial R_0 of less than one can still establish an epidemic after evolutionary rescue. If a pathogen is maintained in a novel host population by sufficient stochastic transmission chains and its mutation rate is not 0, novel mutations can arise that increase R_0 to or above 1. When multiple evolutionary changes are required for rescue, the probability of emergence decreases and becomes much more sensitive to the initial value of R_0. Evolutionary change must occur sufficiently fast to rescue a novel pathogen from extinction and lead to emergence.

Novel pathogens may circulate back into a reservoir population, potentially increasing the length of stochastic transmission chains and contributing to evolutionary rescue and emergence. Assuming that a mutation that increases transmission in the novel host is neutral in the reservoir, the time needed for further adaptive mutations to arise in the pathogen may be extended (Reluga *et al.*, 2007). The process of back-circulation to a reservoir population is conceptually similar to a scenario in which novel hosts vary in epidemiologically-relevant ways, such that some hosts are “higher quality” from the perspective of the pathogen.

Most theoretical models treat host populations as homogeneous. However, this assumption is broken in nature. Recent progress has been made to overcome this disconnect and shed light on how host diversity can limit both the spread and evolution of pathogens.

It is clear that host genetic diversity can successfully limit pathogen spread. This can be explained by the dilution effect (Keesing *et al.*, 2006; Keesing *et al.*, 2010; Ostfeld & Keesing, 2012). The dilution effect, in terms of within-species immune diversity, is suggested to arise because increasing the number of resistant or low-quality individuals in a host population decreases the proportion of susceptible individuals in that population. This can reduce contact rates between free-living pathogens and susceptible hosts (Mitchell *et al.*, 2002; Dobson, 2004; Gandon, 2004; Dennehy *et al.*, 2007; Roscher *et al.*, 2007; Lively, 2010) or increase the likelihood of stochastic fixation of resistance alleles (Quigley *et al.*, 2012; Gokhale *et al.*, 2013), which in turn can reduce R_0. There is much observational support for the role of a dilution effect in multi-species host–pathogen systems [reviewed in Civitello *et al.* (2015)]. Further,

the dilution effect is analogous with the concept of herd immunity in epidemiology (Topley & Wilson, 1923). Herd immunity is based on the idea that the presence and/or proximity of immune or partially-immune individuals reduces the risk of infection for susceptible individuals by suppressing R_0 (Fox *et al.*, 1971; Anderson & May, 1985; John & Samuel, 2000; Stephens, 2008). The public health application of herd immunity (or the dilution effect) relates to the critical vaccination level V_C, which is the proportion of the population that must be vaccinated in order to achieve herd immunity and an $R_0 < 1$ (Anderson & May, 1992; Fine, 1993; Heesterbeek, 2002). Evidence for the ability of herd immunity to successfully limit and in some cases, eradicate, a pathogen, is widespread in the epidemiological literature (Hamer, 1906; Hedrich, 1933; Böttiger, 1987; Reichert *et al.*, 2001; Heymann & Aylward, 2008; Fine *et al.*, 2011).

The dilution effect in these contexts is a *relative* dilution effect – host diversity reduces contact probability between infected and susceptible hosts, which in turn reduces pathogen transmission. Crucially, this is functionally different from an *absolute* dilution effect, where contact probability and transmission is limited instead by spatiotemporal factors. This is a non-trivial distinction, as absolute and relative dilution effects are likely to differ in their ecological and evolutionary consequences. Selection for pathogen escape may be elevated by relative dilution due to more frequent contact with less-suitable hosts; host-pathogen coexistence and coevolution under absolute dilution may be limited by the spatiotemporal availability of host resources instead of host diversity; or stochastic transmission chains required for evolutionary emergence may be more likely to be maintained when less-suitable

hosts are available under relative dilution. The process and patterns of absolute versus relative dilution effects are explored in Chapters 2 and 3, respectively.

While the dilution effect is most often applied to describe why existing pathogens are limited by host diversity, the process is analogous to the idea that diversity reduces stochastic transmission chains discussed in detail above. Extending the evolutionary epidemiology framework of stochastic transmission chains while $R_0 < 1$ to account for host heterogeneity, for example in terms of susceptibility, infectiousness, or contact patterns, has suggested several mechanisms by which host diversity may influence pathogen evolution. When a heterogeneous population consists of hosts that vary in their individual R_0 i.e. super-spreaders (Woolhouse *et al.*, 1997), the probability of emergence is reduced but the epidemic is larger if evolutionary rescue does occur (Lloyd-Smith *et al.*, 2005). This is because heterogeneity can reduce the length and/or frequency of stochastic transmission chains owing to the presence of low-quality hosts, an effect which is not offset by the presence of high-quality hosts. However, in the event that adaptation leads to an $R_0 > 1$, the presence of high-quality hosts increases the size of the resultant epidemic (Lloyd-Smith *et al.*, 2005).

In general, theory predicts that the probability of evolutionary emergence decreases when host diversity increases, but this relationship is thought to be non-linear. Benmayor *et al.* (2009) showed that evolutionary emergence of novel phage Φ2 genotypes peaked at intermediate frequencies of susceptible *P. fluorescens* hosts due to an optimal balance between selection pressure, contact rates, and mutation supply. Chabas *et al.* (2018) extend this using a mechanistic model of pathogen emergence in a diverse host population where hosts differ in how they resist

infection. Their model predicts that evolutionary emergence is maximised at intermediate susceptible host frequencies, and experimentally validate these predictions using hosts populations that vary in CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats; CRISPR-associated) allele diversity challenged with lytic phage. Building on theory discussed throughout this section and in agreement with Benmayor *et al.* (2009), Chabas *et al.* (2018) suggest that this effect of intermediate host heterogeneity is due to two opposing selective forces. While $R_0 < 1$, a higher frequency of susceptible hosts increases the length of stochastic transmission chains that limit pathogen extinction. Simultaneously, increasing host diversity increases the selection pressure for novel variants. If the frequency of susceptible hosts is high due to low host diversity, then pathogen reproduction is high and there is limited selection for evolutionary novelty. In the most extreme scenario, a monoclonal host population, the pathogen can infect all individuals and there is no selection. By contrast, when the frequency of susceptible hosts is low due to high host diversity, transmission chains go extinct too quickly for novel variants to survive and increase R_0 above 1. These two effects, acting in dynamic tension, can consequently maximise evolutionary emergence at intermediate host diversity.

By contrast, a phenomenological approach that makes explicit the contact process between hosts and the network of mutations available to the pathogen suggest that host heterogeneity may actually increase the likelihood of evolutionary emergence (Alexander & Day, 2010), and the evolutionary emergence of phage Φ6 has been shown to increase when advantageous mutations are “nearby” in the mutational network (Burch & Chao, 2000). This approach only accounts for differences in

contact rates, rather than host susceptibility. Despite these differing predictions, the probability of emergence in a heterogeneous host following evolutionary rescue is particularly sensitive to stochastic effects when $R_0 \sim 1$ (Yates *et al.*, 2006; Alexander & Day, 2010).

Experimental tests: bacteria-phage

One of the major challenges to testing theoretical predictions of the effect of host diversity on pathogen spread is that, in almost every host-pathogen system, the genetics of resistance and infectivity are either unknown or not amenable to experimentation. Experiments with bacteria and phage have been used to test some of the theoretical predictions on the role of host diversity in pathogen evolutionary emergence and ecology. Several excellent reviews discuss why and how studies of bacteria-phage have contributed to our understanding of host-pathogen interactions, for example in terms of linking phenotypic with genotypic change, coevolutionary dynamics, and microbial community processes (Bohannan & Lenski, 2000; Buckling & Brockhurst, 2012; Dennehy, 2012; Koskella, 2014; Koskella & Brockhurst, 2014). Of particular relevance to this thesis is the contribution of bacteria-phage studies to our understanding of the evolution of host range, and its relationship to host-pathogen coevolution.

Early work in several different bacteria-phage systems, but predominantly using *Escherichia coli* B, suggested that bacteria-phage coevolution was fundamentally limited [reviewed in Dennehy (2012)]. To explain this, Lenski and Levin (1985) suggested that coevolution ceased once *de novo* modifications to cell surface

receptors arose that prevented phage attachment. Further, they argued that bacteria-phage are locked into “mutational asymmetry”. This arises because phage must bind to a specific surface receptor to be infective. If that receptor changes or is lost, phage must modify the interaction or switch to an entirely new receptor. Lenski and Levin (1985) argued that this is much more mutationally and evolutionarily constrained compared to bacteria, which have multiple routes to receptor modification or loss. In other words, the “supply” of possible mutations to phage is much less than that of bacteria. The potential for long-term coevolution is therefore limited by the ceiling on phage host range evolution imposed by mutational asymmetry.

Although mutational asymmetry is probably a key constraint on bacteria-phage coevolution, the influence of other evolutionary and ecological forces can maintain bacteria-phage coevolution and coexistence. Many bacteria-phage systems show that host range evolution, and bacteria-phage coevolution, are not necessarily as limited as Lenski and Levin (1985) suggest. Prolonged periods of arms race coevolution (Dawkins, 1979), where bacteria evolve to become more broadly resistant and phage evolve to become more broadly infective (Buckling & Rainey, 2002b, 2002a), have been observed in laboratory cultures of several species (Buckling & Rainey, 2002a; Mizoguchi *et al.*, 2003; Wichman *et al.*, 2005; Middelboe *et al.*, 2009; Hall *et al.*, 2011b; Marston *et al.*, 2012; Paez-Espino *et al.*, 2015), including *E. coli* B (Meyer *et al.*, 2012).

While the evidence for arms races between bacteria and phage is strong, there do still seem to be fundamental limits to phage host range expansion. Scanlan *et al.*

(2013) showed that, while Φ2 increased its host range against sympatric and allopatric coevolved *P. fluorescens* under an arms race dynamic, this did not confer broader infectivity against other *P. fluorescens* variants. This suggests that phage host range evolution may still constrained to their specific coevolutionary context. In nature, phage tend toward specialism rather than generalism (Hyman & Abedon, 2010; Koskella & Meaden, 2013; Martiny *et al.*, 2014). Analyses of bacteria-phage interaction networks, which represent and quantify the patterns of infectivity and resistance in a bacteria-phage community (Weitz *et al.*, 2013), show that while phage can indeed occupy all parts of the specialist-generalist spectrum in nature (Flores *et al.*, 2013; Poisot *et al.*, 2013), these networks tend to have a nested structure, with a few common generalist and many rare specialist phage genotypes (Weitz *et al.*, 2013; Martiny *et al.*, 2014). It is interesting to consider that this pattern of “few common, many rare” in bacteria-phage interactions is consistent with more general patterns of community assembly at different ecological levels (Lawton, 1999; Magurran & Henderson, 2003; Verberk *et al.*, 2010).

Aside from mutational asymmetry, what other factors may limit phage host range expansion and favour specialism over generalism? Fitness trade-offs, where the evolution of a broader host range incurs the cost of losing infectivity on the original host, have been shown to limit host range expansion of Φ6 infecting *P. syringae* (Duffy *et al.*, 2006). Further, Hall *et al.* (2011b) show that increasing costs of generalism eventually weaken the arms race dynamic between *P. fluorescens* and Φ2. This can give way to fluctuating selection dynamics (Gandon *et al.*, 2008) and increased specialisation. Fitness trade-offs and the costs of generalism may also scale with the diversity or heterogeneity of the host population, where losing

infectivity on the original host may be more likely to limit phage transmission when susceptible hosts are less available.

Variability in host quality has also been suggested to shape phage host range evolution in terms of optimal foraging. Optimal foraging theory assumes that resources (host cells), are heterogeneous in terms of quality or spatiotemporal availability, and that the costs of increased host range are incurred by lost opportunities to exploit better-quality hosts (MacArthur & Pianka, 1966; Pyke *et al.*, 1977). The more likely an encounter with a lower-quality host, the more advantageous it is to specialise on a higher-quality host – better the devil you know. Optimal foraging also predicts that the degree of specialisation is dependent on host density. Heineman *et al.* (2008) showed that specialist T7 phage were selected for when higher-quality *E. coli* hosts were common, or when hosts significantly varied in quality, and that the degree of specialisation indeed dependent on host density. However, spatial heterogeneity did not lead to increased specialisation in the *P. fluorescens*-Φ2 system (Hesse *et al.*, 2015), which may be due to other selective forces imposed on the host by diverse environmental conditions. Finally, Guyader and Burch (2008) showed that generalist phage did not evolve toward increased specialisation at high host densities, and *vice versa*, as would be expected if optimal foraging shaped phage host range evolution. Together, genetic constraints in terms of mutational asymmetry and fitness trade-offs, rather than ecological constraints in terms of host resource patchiness, are likely to play a larger role in limiting phage host range evolution. The importance of genetic constraints, especially in diverse host populations, has been explored further in bacteria-phage systems where hosts employ a CRISPR-Cas immune system.

Experimental tests – CRISPR-Cas systems

Most recently, there has been increasing interest in how the CRISPR-Cas bacterial immune system can be used to inform a more general understanding of how diversity shape host-pathogen ecology and evolution. CRISPR-Cas immune systems can incorporate short DNA fragments (spacers) of about 30 base pairs (bp) derived from the phage genome into CRISPR loci on the host genome (Horvath *et al.*, 2008). Processed CRISPR transcripts guide Cas immune complexes to identify and cleave the invading phage genome, preventing successful re-infections (Brouns *et al.*, 2008; Marraffini & Sontheimer, 2008; Garneau *et al.*, 2010; Datsenko *et al.*, 2012; Westra *et al.*, 2012b). Because the process of spacer acquisition samples genetic material from the phage population, CRISPR-Cas can rapidly generate high levels of within- and between-host diversity. This has been observed by several studies in both natural and experimental bacteria-phage communities (Andersson & Banfield, 2008; Paez-Espino *et al.*, 2013; Westra *et al.*, 2015; Pyenson & Marraffini, 2020). The diversity generated by CRISPR-Cas is heritable because phage-derived spacers are passed to daughter cells. Although diversity is heritable, spacers at the CRISPR loci can be lost over time (W. Jiang *et al.*, 2013; Chaudhry *et al.*, 2018). Phage can evolve to overcome CRISPR immunity by point mutation in the sequence targeted by the spacer (protospacer) or in the protospacer-adjacent motif (PAM), which flanks the protospacer and functions in self/non self-discrimination (Deveau *et al.*, 2008; Mojica *et al.*, 2009; Semenova *et al.*, 2011; Westra *et al.*, 2013).

The mechanistic features of CRISPR-Cas and phage, as well as their coevolution, make it an ideal model system with which to study the evolutionary ecology of infectious diseases, and especially the role of host diversity. Because the molecular basis of host immunity and phage escape is well-understood and relatively easy to characterise using PCR and Sanger sequencing, phenotypic changes (host resistance, phage infectivity) can be linked to genotypic changes (spacer acquisition, protospacer evolution). Host diversity can be easily experimentally manipulated by mixing bacterial clones with different CRISPR loci, and is also straightforward to measure in downstream analyses. One of the most exciting and innovative applications of CRISPR-Cas to the question of host diversity is that it can used to generate different genetic architectures of host-pathogen resistance-infectivity that directly reflect the core assumptions of many mathematical models of coevolution (Agrawal & Lively, 2002; Lively, 2010).

CRISPR diversity can provide increased resistance by limiting the ability of phage to acquire the mutations needed to overcome CRISPR immunity. Phage evolution can be limited by mutation supply, as discussed above (Lenski & Levin, 1985). Mutation supply in the context of CRISPR occurs because the expansion of a host's spacer repertoire offers multiple mutational routes to phage resistance at little to no fitness cost (Vale *et al.*, 2015), while phage are relatively more constrained by the need to alter targeted protospacers through SNPs or more significant mutations in order to evade CRISPR-Cas interference (Levin *et al.*, 2013; Chabas *et al.*, 2018). Further, full phage infectivity requires mutations in all the protospacers targeted by the host CRISPR array. This is similar to the suggested reason for mutational asymmetry in *P. fluorescens* and phage Φ2, whereby the evolution of host range expansion

requires the stepwise addition of mutations (Hall *et al.*, 2011a; Hall *et al.*, 2011b). This becomes more difficult as individual-and population-level CRISPR diversity increases (Levin *et al.*, 2013; van Houte *et al.*, 2016b), and can drive rapid phage extinction in both *P. aeruginosa* and *S. thermophilus* (van Houte *et al.*, 2016b; Morley *et al.*, 2017; Chabas *et al.*, 2018). Finally, community diversity has been shown to alter the relative fitness of CRISPR-Cas compared to surface modification (SM), with consequences for the virulence of a mixed bacterial infection (Alseth *et al.*, 2019). More specifically, understanding how ecology shapes the selective forces that favour different immune systems has important implications for treating clinical bacterial infections and optimising phage therapy (Westra *et al.*, 2012a; van Houte *et al.*, 2016a; Jackson *et al.*, 2017; Westra *et al.*, 2017).

Generalisations

Although experiments using bacteria-phage as a model system have yielded many novel insights into the role of diversity in host-pathogen dynamics, this approach is inevitably associated with limitations regarding generality. Combining and complimenting it with theory helps to generalise e.g. Chabas *et al.* (2018). Moreover, while bacteria may seem far from eukaryotes, certain core principles are conserved. For example, the way in which CRISPR generates diversity in host populations is conceptually similar to the vertebrate immune system. The major histocompatibility complex (MHC) underpins the adaptive immune response in jawed vertebrates, and particularly so in mammals. Importantly, MHC genes are diverse: they frequently emerge as “hot spots” in genome-wide association studies (GWAS) in humans, and tend to carry on average more SNPs, indels, gene duplications, transposable

element insertions and inversions than the rest of the genome (Chuang & Li, 2004). They are also hot spots for recombination during gamete formation (Paigen & Petkov, 2010). The high relative diversity of MHC genes is generated by higher rates of germline mutagenesis and sexual reproduction. While MHC genes can generate high host diversity from one generation to the next, V(D)J (Variable, Joining, sometimes Diversity) recombination generates immense between- and within-host diversity in a single generation (Early *et al.*, 1980; Rast & Litman, 1994). The V(D)J mechanism can generate highly-diverse repertoires of B and T cell receptors and antibodies (Weinstein *et al.*, 2009; N. Jiang *et al.*, 2011). The close analogies between CRISPR-Cas and vertebrate immune systems as diversity-generating mechanisms (DGMs) (Westra *et al.*, 2017) highlights the wider significance of using bacteria-phage models to understand the role of diversity in host-pathogen systems.

Aims & objectives

The aim of this thesis is to contribute novel understanding of how diversity in host populations affects the evolution and ecology of host-pathogen interactions. In the next three chapters I aim to address the following questions:

1. How do population bottlenecks affect bacteria-phage coexistence?

A population bottleneck is arguably the most frequent way in which population diversity is reduced. They capture multiple natural and anthropogenic causes of diversity reduction and are a recurrent outcome of host-pathogen ecology. Despite their pervasiveness, there has been relatively little consideration of how bottlenecks, as stochastic limitations of diversity, affect host-pathogen dynamics. An objective of

this thesis is to use bacteria-phage to explore if and how bottlenecking limits bacteria-phage coexistence, particularly when hosts are able to generate immune diversity through the CRISPR-Cas system.

2. How does host diversity affect the ability of an evolved pathogen to establish an epidemic?

Host diversity can influence the evolutionary emergence of a novel pathogen, as well as limit the epidemic spread of existing pathogens. All of the work in this area has so far been limited to considering how maladapted pathogens can be “rescued” by evolutionary change, or descriptions of how host diversity affects epidemics. However, insights from host-pathogen coevolution and the evolutionary dynamics of human pathogens tell us that these processes are intimately related and often the same. A key unanswered question is: how is the spread of an evolved pathogen affected by host diversity? An objective of this thesis is to utilise the clear mechanistic understanding of CRISPR-phage coevolution to explore further the link between the evolutionary and ecological effects of host diversity on host-pathogen coexistence.

3. Can coevolution enable bacteria-phage coexistence in the context of a diversity-generating mechanism?

Hosts and pathogens can clearly coexist over long periods of time, and ongoing coevolution often explains this coexistence in many different systems. In bacteria-phage communities where hosts use CRISPR-Cas to defend against phage, the role of coevolution and its dynamic is disputed. Given that CRISPR-Cas is a DGM and thus spontaneously generates host diversity, an objective of this thesis is explore

what kind of coevolutionary dynamic is possible in this context and the molecular changes that explain bacteria-phage coexistence.

Results summary

Chapter 2 explores how population bottlenecking affects the population and evolutionary dynamics between *P. aeruginosa* PA14 and its lytic phage, DMS3vir. Phage went extinct more rapidly, and sensitive host phenotypes invaded the culture, as bottlenecking increased. The data suggested that increased bottlenecking limited phage amplification, which protected the *P. aeruginosa* host population. With some exceptions, this effect was not caused by CRISPR immunity. Controls with a CRISPR knockout strain showed the same phage dynamics. This suggested that the host benefit of bottlenecking is driven by a dilution effect that disproportionately affects phage. Confirming this, experiments where populations of equivalent size were diluted to different degrees showed qualitatively similar phage dynamics. From the host perspective, the dilution effect should protect sensitive hosts, hence the invasion of sensitive cells as bottlenecking increased. Experiments where only the bacterial fraction of the population was bottlenecked showed that CRISPR immunity was maintained. This confirmed that the protective dilution effect was removed, maintaining selection for phage resistance across the gradient of bottleneck size.

Chapter 3 builds on Chapter 2 to further explore the dilution effect in bacteria-phage coevolution. Instead of population bottlenecks, this chapter explicitly examined how the level of CRISPR diversity in the host CRISPR-Cas system affected phage dynamics and evolutionary emergence, as well as host dynamics, and the fitness of

resistant and susceptible host genotypes. A host-pathogen system was engineered where each bacterial host genotype could be infected by only one phage genotype. Experiments using this model system explored how CRISPR diversity impacts the spread of phage when they can overcome a resistance allele, how immune diversity affects the evolution of the phage to increase its host range, and if there was feedback between these processes. These data showed that increasing CRISPR diversity benefits susceptible bacteria via a dilution effect, which limits the spread of the phage. This study suggests that this ecological effect impacts the evolution of novel phage genotypes, which then feeds back into phage population dynamics.

Chapter 4 shifts focus from the mechanisms by which host diversity affects bacteria-phage interactions, and explores the nature of the coevolutionary dynamic between bacteria and phage. It also utilised a different model system – *Streptococcus thermophilus* and its lytic phage 2972. Over 30 days, *S. thermophilus* and phage 2972 were able to coexist for many generations. Phenotypic and genotypic analysis of the bacterial population showed that this coexistence was enabled by a coevolutionary arms race dynamic (ARD) between the host's CRISPR-mediated resistance and phage infectivity. Asymmetry between phage infectivity and host resistance due to limitations on phage mutation supply eventually caused phage extinction. Together, these data show that a bacterial species with a CRISPR system can coexist with phage over long periods because of an ARD.

Contribution of published papers

Chapter 2 - Common J & Westra ER (2019) "CRISPR evolution and bacteriophage persistence in the context of population bottlenecks" *RNA Biology* 16:4, 377-379: This study provides novel insight into how bottlenecking influences bacteria-phage dynamics by emphasizing the role of dilution in bacteria-phage interactions.

Chapter 3 - Common J, Walker-Sünderhauf D, van Houte S, & Westra ER (2020) "Diversity in CRISPR-based immunity protects susceptible genotypes by restricting phage spread and evolution" *Journal of Evolutionary Biology* 33: 1097-1108: This study shows a clear mechanism by which host diversity can limit pathogen spread. Such a mechanism has not previously been shown in any empirical host-pathogen system.

Chapter 3: Common J, Morley D, Westra ER, & van Houte S (2019) "CRISPR-Cas immunity leads to a coevolutionary arms race between *Streptococcus thermophilus* and lytic phage" *Philosophical Transactions of the Royal Society B: Biological Sciences* 374: This study provides novel insight in the way CRISPR-Cas systems shape the population and coevolutionary dynamics of bacteria-phage interactions.

Statement of contributions as co-author

My contributions as co-author to the published papers included in this thesis are as follows. For each, I acknowledge and thank the anonymous reviewers, whose suggestions and criticisms greatly improved the papers.

Chapter 2 - Common J & Westra ER (2019): I was involved in the design of experiments, conducted all experiments and analyses, and wrote the manuscript. ERW conceived of the study, co-designed experiments, and provided critical revisions on the manuscript.

Chapter 3 - Common J, Walker-Sünderhauf D, van Houte S, & Westra ER (2020): I was involved in the design of experiments, conducted all experiments and analyses, and wrote the manuscript. DWS and SVH designed the plasmid used to tag strains with *lacZ*. ERW conceived of the study and co-designed experiments. DWS, SVH and ERW provided critical revisions on the manuscript.

Chapter 4 - Common J, Morley D, Westra ER, & van Houte S (2019): I conducted all statistical and genomic analyses and carried out part of the experimental work. The outcome of these analyses contributed significantly to the direction and narrative of the paper. DM conducted part of the experiments. SVH and ERW conceived the study and designed experiments. All authors contributed to manuscript writing.

Chapter 2: CRISPR evolution and bacteriophage persistence in the context of population bottlenecks

Publication details

The contents of this chapter have been published as the following:

> Common J & Westra ER (2019) “CRISPR evolution and bacteriophage persistence in the context of population bottlenecks” *RNA Biology* 16:4, 377-379 DOI: 10.1080/15476286.2019.1578608

The main text, figures, tables, and supplemental information are those which appear in publication. They have been formatted according to submission requirements and to provide consistency throughout the thesis. Details specific to publication, such as funding details and data accessibility, have been excluded. Author contributions can be found under Statement of contributions as co-author.

Abstract

Population bottlenecks often cause strong reductions in genetic diversity and alter population structure. In the context of host-parasite interactions, bottlenecks could in theory benefit either the host or the pathogen. We predicted that bottlenecking of bacterial populations that evolve CRISPR immunity against bacteriophages (phage) would benefit the pathogen, because CRISPR spacer diversity can rapidly drive phages extinct. To test this, we bottlenecked populations of bacteria and phage, tracking phage persistence and the evolution of bacterial resistance mechanisms. Contrary to our prediction, bottlenecking worked in the advantage of the host. With some exceptions, this effect was not caused by CRISPR immunity. This host benefit is consistent with a dilution effect disproportionately affecting phage. This study provides further insight into how bottlenecking influences bacteria-phage dynamics, the role of dilution in bacteria-phage interactions, and the evolution of host immune systems.

Introduction

Many bacteria encode CRISPR-Cas immune systems [Clustered Regularly Interspaced Short Palindromic Repeats, CRISPR-associated], a RNA-guided mechanism used to defend against phage infection (van Houte *et al.*, 2016a). In response to phage infection, CRISPR-Cas immune systems can incorporate short DNA fragments of about 30 base pairs derived from the phage genome into CRISPR loci on the host genome, termed spacers (Semenova *et al.*, 2011; Swarts *et al.*, 2012). Processed CRISPR transcripts form small RNAs (crRNA) that bind to CRISPR-associated (Cas) proteins which guide the recognition and cleavage of complementary nucleic acid sequences. In the case of phage infection, CRISPR-Cas systems can target and cleave the invading phage genome, preventing successful re-infections (Brouns *et al.*, 2008; Garneau *et al.*, 2010; Datsenko *et al.*, 2012). CRISPR loci in bacterial populations are often diverse (Andersson & Banfield, 2008; Paez-Espino *et al.*, 2013; Westra *et al.*, 2015). High levels of CRISPR spacer diversity naturally evolve in bacterial populations with type I-F CRISPR-Cas systems due to primed spacer acquisition (Richter *et al.*, 2014). Priming relies on a partial match between a pre-existing spacer and the phage genome (Cady & O'Toole, 2011; Westra *et al.*, 2015), and causes an increase in the rate of spacer acquisition. Increasing CRISPR diversity can contribute to synergistic host benefits and cause increasingly rapid phage extinction (Childs *et al.*, 2014; van Houte *et al.*, 2016b; Morley *et al.*, 2017).

The levels of genetic diversity in populations of hosts and parasites are known to play a key role in determining the spread and evolution of infectious diseases. For

example, host genetic diversity can limit the spread of pathogens [reviewed in (Ashby & King, 2015; Jackson *et al.*, 2017)] whereas parasite genetic diversity can increase the ability of parasites to adapt to local hosts (Gandon & Michalakis, 2002). Population bottlenecks - characterised by sudden, often repeated and usually drastic reductions in population size - are common in host-pathogen populations and can strongly reduce their genetic diversity (Hudson *et al.*, 1998; Bohannan & Lenski, 2000). Bottlenecks in host-pathogen systems can be levied by clinical treatment, such as antibiotics (Flanagan *et al.*, 2007); as part of the normal infection cycle of a pathogen [e.g. Lyme Disease (Rego *et al.*, 2014), HIV-1 (Gijsbers *et al.*, 2012), Hepatitis C virus (Bull *et al.*, 2011), reviewed in (Gutiérrez *et al.*, 2012)]; or sudden changes in the abiotic environment, such as soil structure (Torsvik & Øvreås, 2002).

Bottlenecks can influence host-pathogen interactions in several ways. Bottlenecking can benefit the pathogen, as the reduction in host diversity increases the ability of pathogens to adapt to overcome host defenses (Gandon & Michalakis, 2002; Morgan *et al.*, 2005; Morran *et al.*, 2011). Conversely, bottlenecks can limit the spread of parasites by a dilution effect (Keesing *et al.*, 2006; Ostfeld & Keesing, 2012), where low host density reduces contact rate between susceptible individuals [e.g. (Mitchell *et al.*, 2002; Dennehy *et al.*, 2007; Roscher *et al.*, 2007)], or by stochastic fixation of resistance alleles (Quigley *et al.*, 2012; Gokhale *et al.*, 2013). The only experimental test of bottlenecking in a bacteria-bacteriophage (phage) system supported a host benefit, which the authors suggest was due to highly-frequent resistant host phenotypes being more likely to pass through successive bottlenecks and fix in the population (Hesse & Buckling, 2016). Dennehy *et al*. 2007 also found that wildtype *Pseudomonas phaselicola*, which are susceptible to phage,

benefited from a dilution effect when in mixed cultures with a pilus mutant strain that prevents phage attachment. However, bottlenecking would be expected to generally work in the advantage of the parasite when hosts rely on a diversity-generating immune mechanism (Westra *et al.*, 2017) such as CRISPR.

Here we set out to explore the consequences of bottlenecking in the context of CRISPR using *P. aeruginosa* strain PA14, which has a type I-F CRISPR-Cas system and readily evolves CRISPR immunity against its phage DMS3vir (Richter *et al.*, 2014; Westra *et al.*, 2015). Given that the benefits of CRISPR immunity can depend on the levels of population-level spacer diversity in a pre-immunized host population (van Houte *et al.*, 2016b), we predicted that increased bottlenecking of an initially susceptible host population, where CRISPR diversity evolves upon phage infection, could cause a breakdown in host diversity, removing the synergistic benefits of CRISPR, and hence cause an increase in phage persistence.

Results

To understand the effect of bottlenecking on host and phage population dynamics, we infected either the WT PA14 strain or PA14 *csy3::lacZ* strain, which carries an inactive CRISPR-Cas system (hereafter referred to as ΔCRISPR) with phage DMS3vir in liquid media and transferred daily into fresh medium while manipulating the bottleneck size by varying the dilution factor at each daily transfer from 10^{-2} (weakest bottleneck) to 10^{-9} (strongest bottleneck). Ancestral WT PA14 carries no spacers targeting DMS3vir, so is sensitive to infection.

Surprisingly, phage went (close to) extinct in all treatments by 5 days post-infection (d.p.i.), irrespective of the bottleneck strength (**Fig. 2.1**), and importantly CRISPR background had no significant effect on phage titres ($F_{1,334} = 2.12$, $p = 0.15$), although there appeared to be a transient effect at the 10^{-4} and 10^{-5} bottleneck treatments (see Discussion). Phage extinction risks were also almost identical between these host genetic backgrounds for all bottlenecks (**Table A2.1**). Although there was a significant effect of CRISPR background on host titres ($F_{1,432} = 143.2$, $p < 0.0001$), this was likely due to the transiently higher average host densities at 2 d.p.i. in all treatments (**Fig. 2.2**) (see Discussion). Adjusted R^2 comparisons of nested models showed that the effect of CRISPR background explained only 10% more variance than a model excluding the CRISPR background interaction. Otherwise, host dynamics within comparable treatments were similar in CRISPR and ΔCRISPR backgrounds. These data therefore suggest that CRISPR-Cas systems overall have a negligible and transient impact on the short-term phage population dynamics under the bottlenecking regimes that were explored here.

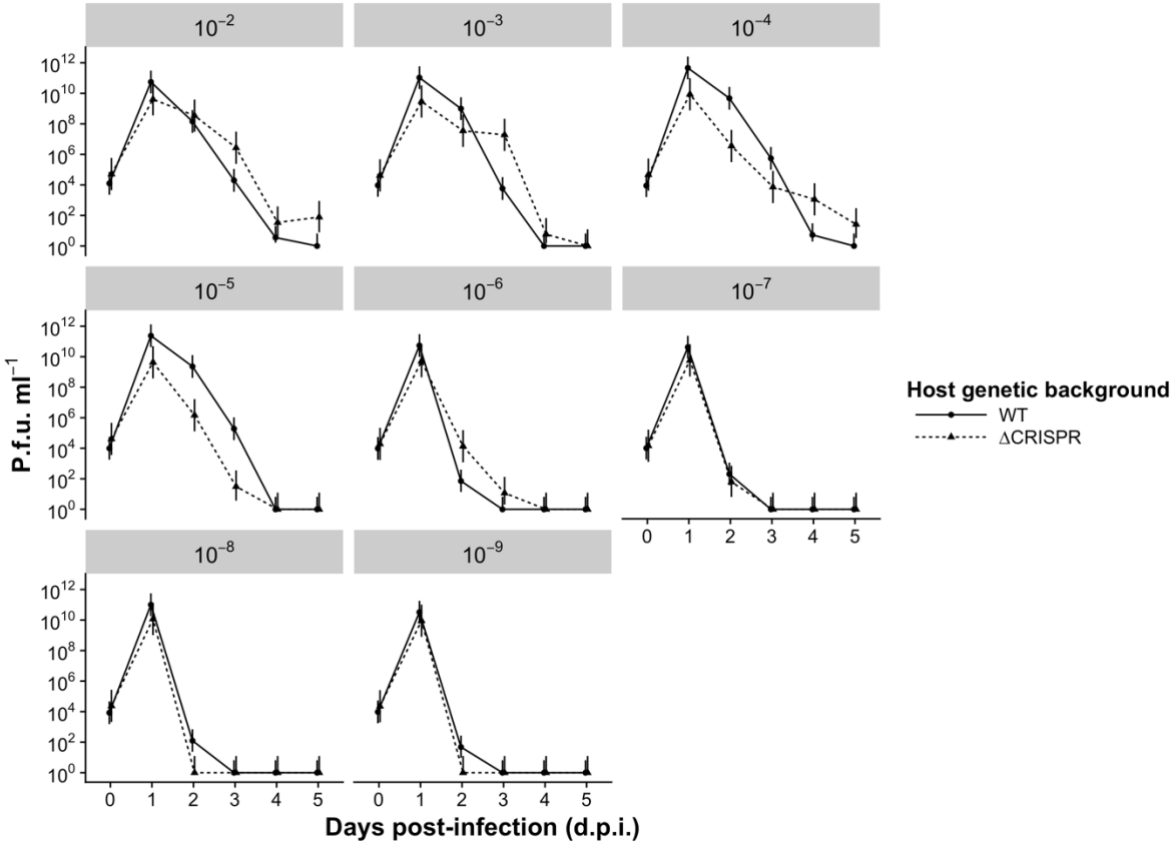


Figure 2.1 Phage population dynamics when both host and phage are bottlenecked

Mean plaque-forming units (p.f.u.) ml–1 for the WT host (i.e. encoding a functional CRISPR-Cas immune system) and the ΔCRISPR strain are shown for different bottleneck treatments (ranging from 10^{-2}-10^{-9} dilutions at each transfer, as indicated above each panel). The detection limit is 200 p.f.u. ml^{-1}. Error bars correspond to 95% confidence intervals (CIs). N = 6 for all treatments.

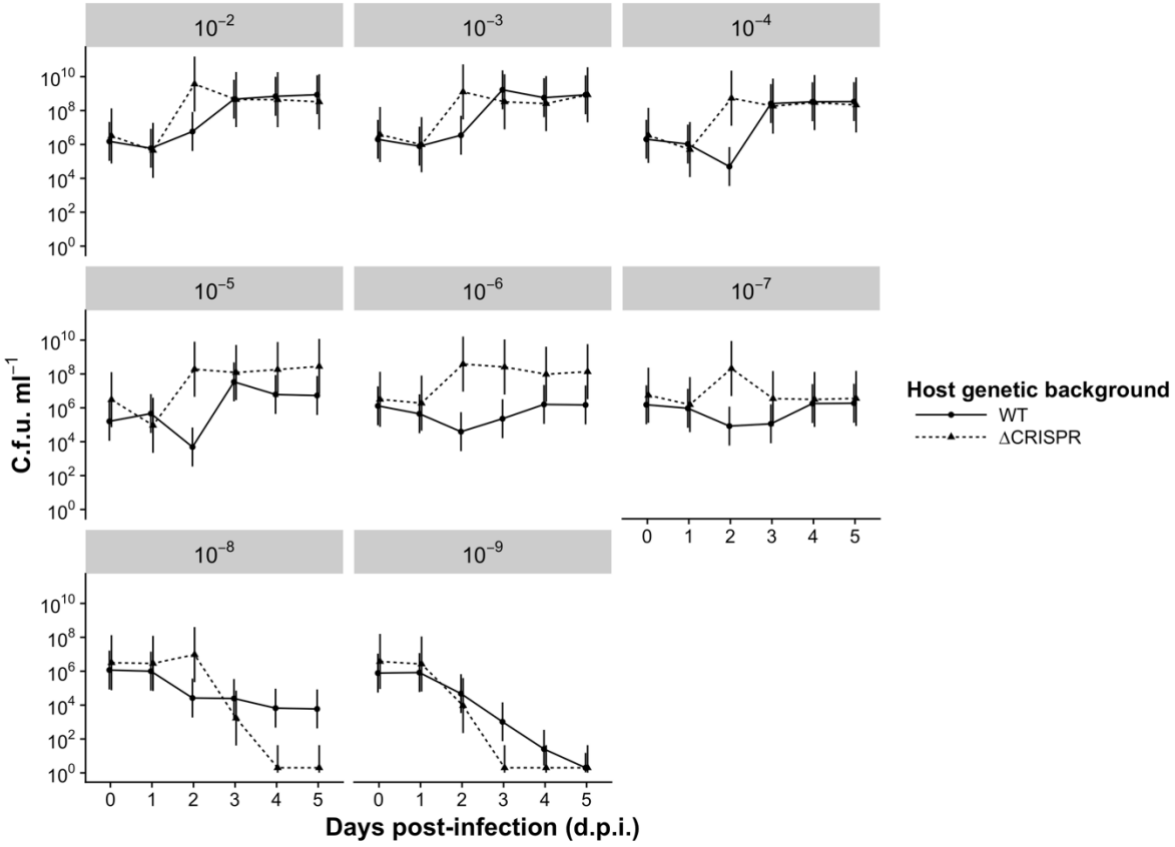


Figure 2.2 Host population dynamics when both host and phage are bottlenecked

Mean colony-forming units (c.f.u.) ml^{-1} for the WT host (i.e. encoding a functional CRISPR-Cas immune system) and the ΔCRISPR strain are shown for different bottleneck treatments (ranging from 10^{-2}-10^{-9} dilutions at each transfer, as indicated above each panel). The detection limit is 200 c.f.u. ml^{-1}. Error bars correspond to 95% confidence intervals (CIs). N = 6 for all treatments.

Although the presence or absence of a functional CRISPR-Cas system did not have a strong impact on the phage population dynamics, the bottlenecking regime itself did have a clear impact with stronger bottlenecks being associated with a more rapid decline in phage titers. This is supported by a control without phage, which shows that host densities remain relatively stable through all but the strongest bottleneck treatments (**Figure A2.1**). We also ruled out any abiotic effects of fresh media that may have disproportionately impacted phage survival [(Jończyk *et al.*, 2011); **Figure A2.2**]. We hypothesized this might be caused by a dilution effect, i.e. a decrease in host densities resulting in reduced parasite spread. To explicitly test this, we first selected two bottleneck treatments (10^{-4} and 10^{-6}) that showed clear differences in phage titers over time. We then set up an experiment as described above with WT PA14 and DMS3vir, and at 1 d.p.i. diluted the culture 100-fold and transferred daily 0.6µl into either 6ml or 600ml of fresh media, generating either 10^{-4} or 10^{-6} dilutions while maintaining the same degree of bottlenecking on the population. We found that the small and large dilutions of bacterial populations led to a qualitatively similar phage population dynamics as those associated with the 10^{-4} and 10^{-6} bottleneck treatments, respectively (**Fig. 2.3**). Phage persisted in most small dilution replicates until 4 d.p.i, while in the large dilution it was driven extinct in all but one replicate by 3 d.p.i. Host persisted for the duration of the experiment in the small dilution treatment, but were undetectable in 5 out of 6 large dilution replicates by 3 d.p.i., and in 3 out of 6 replicates of the 10^{-6} bottleneck treatment, also by 3 d.p.i. These data therefore support the conclusion that phage population dynamics observed in the bottleneck experiment were driven primarily by dilution of the phage. Interestingly, phage titre did not covary significantly with host density in the bottleneck ($F_{68,247} = 1.06$, $p = 0.37$) or dilution experiment ($F_{13,98} = 0.43$, $p = 0.96$), which suggests that phage

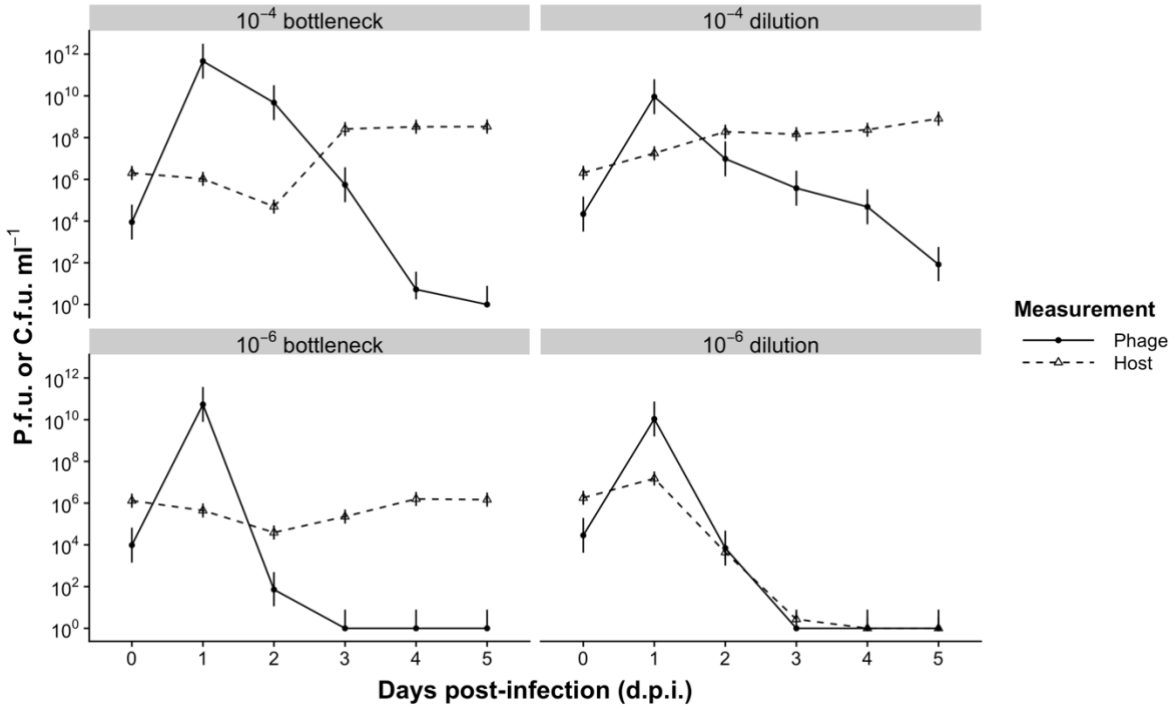


Figure 2.3 Phage and host dynamics when both WT host and phage are diluted or bottlenecked

A 10^{-4} or 10^{-6} bottleneck corresponds to either 0.6µl or 0.006µl culture, respectively, in 6ml of fresh media at each transfer. A 10^{-4} dilution corresponds to 0.6µl culture in 6 ml at each transfer, and a 10^{-6} dilution corresponds to 0.6µl culture in 600 ml at each transfer (see Materials & Methods). Mean plaque-forming units (p.f.u.) ml^{-1} and colony-forming units (c.f.u.) for the dilution and comparable bottleneck treatments are shown. The detection limit is 200 p.f.u. or c.f.u. ml^{-1}. Error bars correspond to 95% confidence intervals (CIs). N = 6 for all treatments.

dynamics are not correlated with those of the host, and consequently are likely more negatively affected by bottlenecking than the host.

Given the observed effect of bottlenecking on the phage population dynamics, we investigated whether this could have knock-on effects for the evolution of host resistance to the phage, and the mechanistic basis for resistance evolution. For consistency with previous studies (Westra *et al.*, 2015), we examined the phage resistance phenotypes at 3 d.p.i. Hosts with CRISPR immunity were most frequent in the 10^{-2} – 10^{-5} bottleneck treatments, which all contained detectable levels of phage at 3 d.p.i. However, bottlenecks of 10^{-6} – 10^{-9} contained mostly sensitive hosts (**Fig. 2.4A** & **Table A2.2**), which we hypothesized was due to the lower levels of phage in those treatments. A qualitatively similar pattern in the evolution of immune phenotypes was observed in the dilution experiment (**Figure A2.3** & **Table A2.4**). This shift from CRISPR immunity to sensitive bacteria was therefore potentially due to relaxed selection for host immunity as a result of phage dilution. To test this idea, we performed a similar experiment where we bottlenecked cultures (both WT and the ΔCRISPR strain) in the same way as described above but with a fixed dilution of 10^{-2} for the phage. This experimental design therefore results in bottlenecking and dilution of the hosts while maintaining a similar phage pressure across the bottleneck treatments. Under these conditions, CRISPR immunity was maintained across the range of treatments and there was no invasion of sensitives (**Fig. 2.4B** & **Table A2.3**). There was again no significant covariance between phage titre and host density in this control ($F_{65,250} = 1.23$, $p = 0.14$), so that even when phage levels were maintained through each bottleneck, host and phage dynamics were not correlated. Although sensitive invasion did not occur in a ΔCRISPR background when phage

were supplemented, intriguingly there was also no sensitive invasion when both host and phage were bottlenecked (**Figure A2.4**).

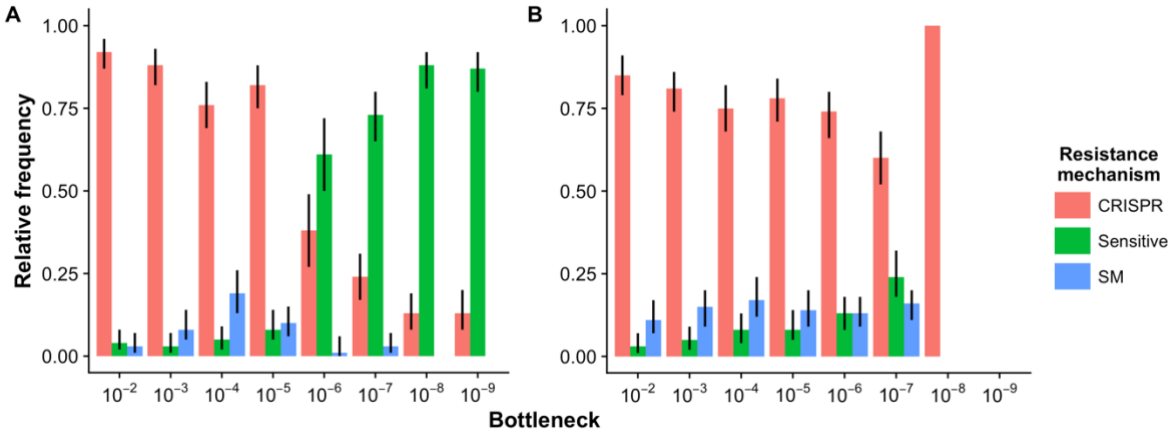


Figure 2.4 Evolution of host resistance of the WT strain

Relative frequencies of bacterial clones with CRISPR immunity, surface mutation (SM) resistance, or sensitive phenotypes, at 3 days post-infection when **A**) both host and phage were bottlenecked by dilution into fresh medium as indicated on the x-axis s and **B**) host was bottlenecked by dilution into fresh medium as indicated on the x-axis and phage was bottlenecked at a fixed 10^{-2} dilution at each transfer. Error bars correspond to 95% confidence intervals (CIs) of the mean. N=72 for the 10^{-6} treatment in panel A, and N=48 for the 10^{-8} treatment in panel B. No bacterial cells were recovered from any replicate in the 10^{-9} treatment in panel B. N =144 for all other treatments.

Discussion

Population bottlenecks often cause strong reductions in genetic diversity (Hudson *et al.*, 1998). In the context of host-parasite interactions, bottlenecks can in theory benefit either the host (Quigley *et al.*, 2012; Gokhale *et al.*, 2013) or the pathogen (Gandon & Michalakis, 2002; Morgan *et al.*, 2005; Morran *et al.*, 2011) by affecting host-pathogen coevolution. Bottlenecks may also benefit the host through a dilution effect (Keesing *et al.*, 2006; Ostfeld & Keesing, 2012). While there is limited empirical support for a dilution effect in a bacteria-phage system (Dennehy *et al.*, 2007), and that bottlenecking benefits bacterial hosts (Hesse & Buckling, 2016), we predicted that bottlenecking of bacterial populations that evolve CRISPR immunity against phages would benefit the pathogen (Westra *et al.*, 2017), because CRISPR spacer diversity can rapidly drive phages extinct (van Houte *et al.*, 2016b).

In contrast to our predictions, bottlenecking did not provide a clear advantage to the phage. Instead, phage were always driven extinct by 5 d.p.i. irrespective of the degree of population bottlenecking. Surprisingly, this effect was not CRISPR-specific but seemed to be driven instead by bottlenecking *per se*, as suggested by similar phage dynamics in the context of a CRISPR-knockout strain. Exploring the potential role of a dilution effect through bottlenecking, we found that dilution alone is sufficient to drive phage extinct in a CRISPR background. In support of this, phage density was uncoupled from host density in all experiments and phage extinction was more rapid after 10^{-6} bottlenecks in cultures with surviving host. Both findings are consistent with dilution disproportionately affecting phage. This biased effect on phage may be related to the different resource requirements of bacteria and phage.

After passing through even very small bottlenecks, bacterial cells can still replicate if environmental resources are available. By contrast, phage cannot replicate independent of hosts - they require a sufficient number and density of host cells to recover after a bottleneck has been applied. Such an imbalance in resource dependence between bacteria and phage, coupled with a reduction in numbers of individual hosts, might explain why phage were more susceptible to dilution. While dilution of hosts can suppress phage epidemics, this will quantitatively depend on phage life history traits, such as adsorption rates, latent period and burst size of the phages. It would be interesting therefore to further explore how these factors affect phage persistence and evolution in the context of host dilution.

From the perspective of the bacterial host, a dilution effect relaxed selection for CRISPR immunity and allowed sensitive hosts to invade, particularly in the stronger bottleneck treatments of 10^{-6} and above. However, we did not see a similar invasion of sensitives in the ΔCRISPR background. It is unclear why sensitives did not invade, but we speculate that it might be attributable to either hosts with CRISPR immunity losing their resistance (Chaudhry *et al.*, 2018; Weissman *et al.*, 2018a), or the CRISPR fraction of the host population 'protecting' sensitive hosts at the start of an infection (Payne *et al.*, 2018), allowing them to invade after already dilute phage have been cleared. When phage were supplemented, sensitives could not invade and CRISPR immunity was most relatively frequent.

Finally, although our results are generally consistent with a dilution being the most important determinant of the phage population dynamics, we did detect some transient effects which are likely driven by CRISPR immunity. First, transiently higher

average host densities at 2 d.p.i. in all treatments (**Fig. 2.2**) may be due to a difference in how quickly hosts with CRISPR immunity versus surface modification (SM) invade the culture. Second, in the experiment where only the host was bottlenecked (resulting in high levels of CRISPR resistance evolution at all treatments), phage titres were statistically similar across bottleneck treatments at all timepoints for the WT bacteria with functional CRISPR-Cas (**Figure A2.5A**), but not for the ΔCRISPR bacteria (**Figure A2.5B**). Third, in the experiment where the whole culture was bottlenecked, phage titres in the 10^{-4} and 10^{-5} treatments were consistently significantly higher at most timepoints compared to the ΔCRISPR background (**Fig. 2.1**), supported by a one-way ANOVA of the effect of host background on phage titre across the whole 6-point time series ($F_{1,36} = 8.92$, $p < 0.01$). These phage titres were maintained despite 10,000- and 100,000-fold reductions of phage immediately after the bottlenecking had been applied, suggesting that phage were still able to successfully amplify on sensitive hosts. High phage titres in these treatments occurred in the context of high relative frequency of CRISPR immunity in the host, raising the possibility of an interaction between CRISPR, host density, and bottlenecking of a certain degree that appears to temporarily favour phage. This suggests that within a certain range of bottleneck treatments CRISPR defenses may become less effective in driving phages extinct, either because of population dynamics or evolutionary effects, or both.

Altogether, this study provides further insight into how bottlenecking influences bacteria-phage dynamics and the evolution of host resistance. Bottlenecking does impact the ability of phage to coexist with the host, but this is generally independent of whether or not the host has a functional CRISPR-Cas immune system. Instead,

our findings support the role of a dilution effect caused by bottlenecking that disproportionately affects phage titre and survival. However, we also find some CRISPR-specific effects under certain conditions that may be related to a complex interaction between host immune systems and host-pathogen dynamics.

Materials & Methods

Bacterial strains & phage

The bacterial strains and phages used in this study have all been described previously. Evolution experiments were carried out using *Pseudomonas aeruginosa* UCBPP-PA14 (WT) and phage DMS3vir (Cady *et al.*, 2012). We used *P. aeuruginosa* UCBPP-PA14 *csy3::lacZ* (Zegans *et al.*, 2009) in CRISPR-negative control experiments and streak assays, and DMS3vir-AcrF1 (van Houte *et al.*, 2016b) in streak assays.

Bottleneck experiments

All experiments were established in glass vials by inoculating 6ml of M9 minimal media (supplemented with 0.2% glucose) with ~$1x10^6$ c.f.u. (colony-forming units) from an overnight culture of WT *P. aeruginosa* PA14. Approximately $1x10^4$ p.f.u. (plaque-forming units) of DMS3vir was added to each glass vial. A phage-negative control was established similarly but without the addition of phage. A CRISPR-negative control was also established by inoculating 6ml of M9 minimal media with ~$1x10^6$ c.f.u. of *P. aeruginosa* PA14 *csy3::lacZ* and ~$1x10^4$ of DMS3vir. At this point (0 d.p.i.), 180µl of culture was taken from each vial and phage was extracted using chloroform. To determine phage titres, extracted phage was then serially diluted eight times in 96-well plates, and 5µl of each phage dilution was spotted on a top lawn of *P. aeruginosa* PA14 *csy3::lacZ*. The detection limit of phage spot assays is 10^2 p.f.u. ml^{-1}. Samples of culture were serially diluted in M9 minimal

media, plated on LB agar and incubated overnight at 37˚C. C.f.u.s were counted to determine host densities. The vials were then incubated at 37˚C while shaking at 180rpm.

After 24hrs of growth (1 d.p.i.) the phage and host sampling protocol, described above, was repeated. To generate bottleneck treatments, samples of culture were then serially diluted in M9 minimal media from 10^{-1}, 10^{-2}, …10^{-7}. 60µl from undiluted culture and each dilution were then used to inoculate 6ml of fresh M9 minimal media. This method of dilution and transfer gives eight bottleneck treatments of 10^{-2}, 10^{-3}, …10^{-9}. To ensure only the host was bottlenecked in the phage-supplemented control, we added 60µl of chloroform-extracted phage from the corresponding replicate into each inoculated microcosm. Each bottleneck was performed in six independent replicates (N = 6). Host and phage sampling, as well as the bottlenecking procedure in each treatment, was repeated 24hrs after growth in fresh medium until 5 d.p.i. Samples of the culture were taken prior to bottlenecking.

Dilution experiment

We explicitly tested the effect of culture dilution on population dynamics and immunity evolution. Twelve glass vials with 6ml M9 minimal media were established with WT *P. aeruginosa* PA14 and DMS3vir as described above, and the host and phage were assayed. We then established two dilution treatments that corresponded to the 10^{-4} and 10^{-6} bottleneck treatments, treatments which showed clear differences in phage titers over time. At 1 d.p.i., all twelve cultures were serially diluted. Six vials (N = 6) of 6ml M9 minimal media were inoculated with 60µl of the

10^{-2} dilution (0.0006ml/6ml = $1x10^{-4}$). Six bottles (*N* = 6) of 600ml M9 minimal media were also inoculated with 60µl of the 10^{-2} dilution (0.0006ml/600ml = $1x10^{-6}$). Sampling (as described above) and dilution were repeated 24hrs after growth in fresh medium until 5 d.p.i. Samples of the culture were taken prior to dilution.

Determining host immune phenotype

Bacterial immunity against ancestral phage was determined at 3 d.p.i, using two independent assays as described in Westra *et al.* (2015). First, bacteria were plated on LB-agar, and 24 randomly-selected individual clones per replicate were streaked through either DMS3vir or DMS3vir-AcrIF1, which encodes an anti-CRISPR gene. Clones sensitive to both phage genotypes were scored as 'sensitive'; those resistant to the DMS3vir but sensitive to DMS3vir-AcrIF1 were scored as 'CRISPR resistant'; and those resistant to both were scored as 'surface mutant (SM)'. Second, each clone was also grown for 24hrs in M9 media in the presence or absence of DMS3vir, and the OD_{600} was measured. Cultures that had a reduced growth rate were scored as sensitive.

Statistical analysis

All statistical analyses were conducted in R (R Core Team, 2018). We took a nested Generalized Linear Model (GLM) approach, and model selection was performed using Aikaike's Information Criteria (AIC) (Akaike, 1973; Burnham & Anderson, 2004). Models of phage titre and host densities used log-transformed residuals to fit the assumption of normality, and coefficients were back-transformed prior to

presentation. Models of phenotype relative frequency used a binomial family with a logit link, and logit coefficients were back-transformed to probability values. We used a Cox proportional hazards model to assess the effect of bottleneck size on phage survivorship over the course of an experiment, with hazard ratio coefficients expressing the relative risk of phage extinction over time. ‘Bottleneck’ was treated as a discrete variable in all models. The package ggplot2 was used to generate graphics (Wickham, 2009).

Acknowledgements

We would like to thank Elze Hesse for advice on survival analyses and comments on the draft of the paper.

Chapter 3: Diversity in CRISPR-based immunity protects susceptible genotypes by restricting phage spread and evolution

Publication details

The contents of this chapter have been published as the following:

> Common J, Walker-Sünderhauf D, van Houte S & Westra ER (2020) "Diversity in CRISPR-based immunity protects susceptible genotypes by restricting phage spread and evolution" *Journal of Evolutionary Biology* 33: 1097-1108 DOI: 10.1111/jeb.13638

The main text, figures, tables, and supplemental information are those which appear in publication. They have been formatted according to submission requirements and to provide consistency throughout the thesis. Details specific to publication, such as funding details and data accessibility, have been excluded. Author contributions can be found under *Statement of contributions as co-author*.

Abstract

Diversity in host resistance often associates with reduced pathogen spread. This may result from ecological and evolutionary processes, likely with feedback between them. Theory and experiments on bacteria-phage interactions have shown that genetic diversity of the bacterial adaptive immune system can limit phage evolution to overcome resistance. Using the CRISPR-Cas bacterial immune system and lytic phage, we engineered a host-pathogen system where each bacterial host genotype could be infected by only one phage genotype. With this model system, we explored how CRISPR diversity impacts the spread of phage when they can overcome a resistance allele, how immune diversity affects the evolution of the phage to increase its host range, and if there was feedback between these processes. We show that increasing CRISPR diversity benefits susceptible bacteria via a dilution effect, which limits the spread of the phage. We suggest that this ecological effect impacts the evolution of novel phage genotypes, which then feeds back into phage population dynamics.

Introduction

Genetic diversity is a key determinant of the ecology and evolution of host-pathogen systems. Various studies of wild organisms have shown that genetic diversity within a host species often affects pathogen prevalence. The importance of diversity for disease prevalence in wild populations has been observed in numerous species, for example: cheetahs (O'Brien *et al.*, 1985); Italian agile frogs (Pearman & Garner, 2005); crimson rosella parrots (Eastwood *et al.*, 2017); inbred black-footed ferret populations (Thorne & Williams, 1988); inbred California sea lions (Acevedo-Whitehouse *et al.*, 2003); young island populations of deer mice (Meagher, 1999); and Galapagos hawks (Whiteman *et al.*, 2007). The importance of diversity for limiting disease in agricultural contexts has long been recognised (Elton, 1958), for example in rice (Zhu *et al.*, 2000) and hybridising populations of honeybees (López-Uribe *et al.*, 2017). In laboratory environments, more genetically diverse populations of *Daphnia magna* are more resistant to parasites (Altermatt & Ebert, 2008), an effect that depends on the genetic architecture of resistance (Luijckx *et al.*, 2013). In microbial systems, *Pseudomonas aeruginosa* PA14 and *Streptococcus thermophilus* with diverse immunity alleles were shown to be more resistant against lytic bacteriophage (van Houte *et al.*, 2016b; Morley *et al.*, 2017). Two recent meta-analyses have shown that host diversity is a robust defence against pathogens both in agricultural (Reiss & Drinkwater, 2018) and natural populations (Ekroth *et al.*, 2019).

The suggested reasons for the effect of host diversity on pathogen spread can be broadly divided into ecological and evolutionary effects. The ecological effect of

diversity may manifest through a dilution effect (Ostfeld & Keesing, 2012; Civitello *et al.*, 2015). The dilution effect, in terms of within-species immune diversity, is suggested to arise because increasing the number of resistant or low-quality individuals in a host population decreases the proportion of susceptible individuals in that population. This can reduce contact rates between free-living pathogens and susceptible hosts, which in turn limits the basic reproduction number of the pathogen (R_0) (Dobson, 2004; Gandon, 2004; Lively, 2010). There is much observational support for the role of a dilution effect in multi-species host-pathogen systems [reviewed in Civitello *et al.*, (2015)], and some experimental work has suggested that dilution of susceptible host genotypes can limit pathogen spread in single host species systems (Dennehy *et al.*, 2007; Common & Westra, 2019). Evolutionary effects of diversity may manifest through limitation of the evolutionary emergence of novel pathogen genotypes (Sasaki, 2000; Ohtsuki & Sasaki, 2006). Importantly though, the evolutionary and ecological effects of host diversity are interdependent as the basic reproductive value of a pathogen will influence its ability to evolve to overcome host resistance (Antia *et al.*, 2003). Because increasing host diversity dilutes susceptible hosts, pathogens may reach smaller population sizes and generate fewer novel variants on which selection can act (Antia *et al.*, 2003; Dennehy *et al.*, 2006; Yates *et al.*, 2006). Simultaneously, increasing host diversity can increase selection for novel variants. These two effects can maximise evolutionary emergence at intermediate host diversity (Benmayor *et al.*, 2009; Chabas *et al.*, 2018).

The interaction between lytic bacteriophage (phage) and the bacterial CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats; CRISPR-associated)

immune system is a tractable model system to study the evolutionary ecology of infectious diseases, including the role of host diversity (van Houte *et al.*, 2016b; Westra *et al.*, 2017; Chabas *et al.*, 2018). CRISPR-Cas immune systems can incorporate short DNA fragments (spacers) of about 30 base pairs derived from the phage genome into CRISPR loci on the host genome (Horvath *et al.*, 2008). Processed CRISPR transcripts guide Cas immune complexes to identify and cleave the invading phage genome, preventing successful re-infections (Brouns *et al.*, 2008; Marraffini & Sontheimer, 2008; Garneau *et al.*, 2010; Datsenko *et al.*, 2012; Westra *et al.*, 2012b). In turn, phage can evolve to overcome CRISPR immunity by point mutation in the sequence targeted by the spacer (protospacer) or in the protospacer-adjacent motif (PAM), which flanks the protospacer and functions in self/non-self discrimination (Deveau *et al.*, 2008; Mojica *et al.*, 2009; Semenova *et al.*, 2011; Westra *et al.*, 2013). Phage evolution to overcome CRISPR immunity can lead to CRISPR-phage coevolution (Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015; Sun *et al.*, 2016; Common *et al.*, 2019). CRISPR loci in both natural and experimental populations can be highly diverse (Andersson & Banfield, 2008; Paez-Espino *et al.*, 2013; Westra *et al.*, 2015; Common *et al.*, 2019), due to different bacteria in the population acquiring different spacers (Westra *et al.*, 2017). Diversity has important implications for the coevolutionary interaction, as CRISPR diversity can provide increased resistance by limiting the ability of phage to acquire the mutations needed to overcome CRISPR immunity (van Houte *et al.*, 2016b; Morley *et al.*, 2017; Common *et al.*, 2019). Phage mutation is limited by mutation supply (Lenski & Levin, 1985; Levin *et al.*, 2013; Chabas *et al.*, 2018), and full phage infectivity requires mutations in all the protospacers targeted by the host CRISPR array. This becomes more difficult as individual-and population-level CRISPR diversity increases (Levin *et*

al., 2013; van Houte *et al.*, 2016b), and can drive rapid phage extinction (van Houte *et al.*, 2016b; Morley *et al.*, 2017; Chabas *et al.*, 2018).

Apart from this evolutionary effect, theory predicts that even if a phage mutant evolved that can overcome one CRISPR resistance allele in the population, its ability to amplify will be reduced in a more diverse host population (Lively, 2010). The resulting smaller phage population sizes are in turn predicted to reduce the ability of the phage to evolve to overcome other CRISPR resistance alleles in the population (Antia *et al.*, 2003; Chabas *et al.*, 2018). However, these predictions remain untested. We therefore set out to explicitly test how host diversity could limit pathogen population size through a dilution effect, if and how this affects phage evolution, and the degree of interdependence between these processes. Using the bacteria *Pseudomonas aeruginosa* and its lytic phage DMS3vir, we performed an experiment where we manipulated the degree of CRISPR diversity in the host population by mixing bacterial genotypes that each carried a different CRISPR spacer. We then infected these host populations with a phage that was pre-evolved to infect only one host genotype. We then tracked changes in host fitness, and phage population dynamics and evolution, over 3 days.

Results

To explore how population-level immune diversity would influence the population dynamics and evolution of an infective phage and its susceptible host genotype, we first developed a library of 24 *P. aeruginosa* PA14 bacteriophage-insensitive mutants (BIMs), and a corresponding library of 24 DMS3vir mutants, each of which was pre-evolved to infect only one of the 24 host BIMs (escape phage; **Fig. 3.1A**). We then set up an experiment where we mixed 1, 3, 6, 12 or 24 BIMs, and inoculated each mixture with ~10^6 pfu ml^{-1} of a single escape phage (**Fig. 3.1B**). Because this experimental design increases the proportion of resistant bacterial hosts while decreasing the proportion of the susceptible host (i.e. the one that could be infected by the escape phage) (**Figure A3.1**), it enables us to explicitly test how this dilution effect is related to the benefits of host CRISPR allele diversity. The susceptible clone always carried a *lacZ* reporter gene, so we could follow its population dynamics and competitive performance during the experiment (**Fig. 3.1B**). *P. aeruginosa* Δ*pilA*, an isogenic mutant which does not have CRISPR immunity and resists phage infection via surface receptor modification, was added to each host mixture (**Fig. 3.1B**). *P. aeruginosa* Δ*pilA* was resistant to all phage, and allowed us to compare the fitness of the CRISPR immune populations across the CRISPR allele diversity levels. We also included 1- and 24-clone treatments inoculated with ancestral phage to which the whole bacterial population was resistant. These treatments allowed us to examine bacteria and phage dynamics at the two extremes of host diversity we tested in host populations with no pre-existing susceptibility. We then compared those dynamics with host mixtures that included a susceptible clone (van Houte *et al.*, 2016b)(**Fig. 3.1B**). We then monitored population dynamics and evolution of the phage, of the

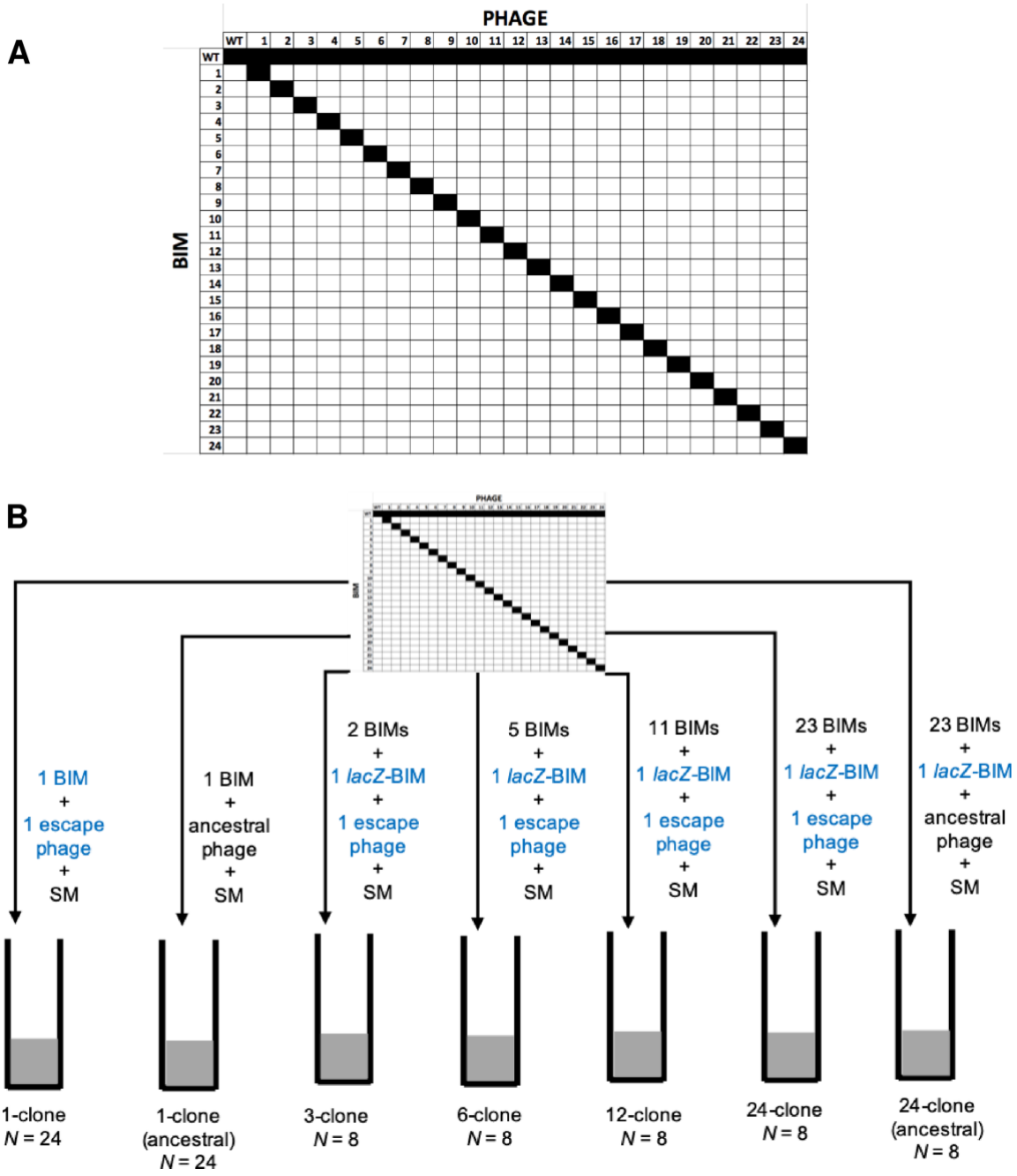


Figure 3.1 Experimental design

(**A**) Infectivity matrix of the library of bacteriophage-insensitive mutants (BIMs) and escape phages. The identity of BIMs and escape phage (1-24) are shown in the first row and column, respectively. Black squares represent infectivity, white indicates no infectivity. Infectivity of the wild-type DMS3vir is shown, as well as the infectivity of

(Figure 3.1 caption contd.)

each escape phage on wild-type *P. aeruginosa* PA14. (**B**) Design of the co-culture experiment. The library of BIMs and escape phage is shown at the top. For each CRISPR allele diversity treatment, the relevant numbers of BIMs and escape phage were taken from the library and used to start the experiment, as shown by the arrows. The number of BIMs is shown next to the arrows. Blue text indicates that the BIM was engineered to carry a *lacZ* reporter gene, and that it was susceptible to the escape phage. Black text indicates that the BIMs were resistant to the phage. The number of replicates (*N*) is given underneath each treatment. SM represents the surface mutant strain PA14 Δ*pilA*.

clones with CRISPR immunity, and the susceptible fraction of the CRISPR population over 3 days.

We hypothesised that if CRISPR diversity limits phage epidemics, then phage density would be negatively correlated with CRISPR diversity. Phage densities decreased as CRISPR diversity increased ($F_{6,342}$ = 48.30, $p < 2.2 \times 10^{-16}$; **Fig. 3.2**), phage population sizes decreased over the course of the experiment (regression of ln plaque-forming units [pfu] ml^{-1} over days post-infection [dpi]: β [95% CI] = -1.32 [-1.74, -0.94], $t(2)_{344}$ = -6.40, $p = 5.06 \times 10^{-10}$), and phage declined faster with increasing CRISPR diversity (interaction between CRISPR diversity and days post-infection: $F_{6,314}$ = 24.10, $p < 2.2 \times 10^{-16}$; **Fig. 3.2A & B**). This is consistent with an effect of CRISPR diversity that protects the bacterial population from phage. Further, phage titre in the 1-clone treatment infected with ancestral phage ["1-clone (ancestral)" in **Fig. 3.2B**] is statistically similar to that observed in the 1-clone treatment infected with an escape phage ["Intercept" in **Fig. 3.2B**], which is in line with previous data showing that monocultures composed of hosts with CRISPR immunity allow phage persistence due to rapid evolution of escape phage that overcome the CRISPR resistance allele (van Houte *et al.*, 2016b). The evolution of escape phage in our monoculture treatments likely caused the observed fluctuations in the density of hosts with CRISPR immunity (**Figure A3.1**). When we explored the model further, we found that phage titres in the 24-clone (ancestral phage) treatment were lower than the 24-clone treatment infected with escape phage (difference in ln pfu ml^{-1}: β [95% CI] = -2.40 [-4.67, -0.33], $t(2)_{342}$ = -2.25, p = 0.03; **Fig. 3.2B**). This is consistent with a modest epidemic being able to establish by replicating on the

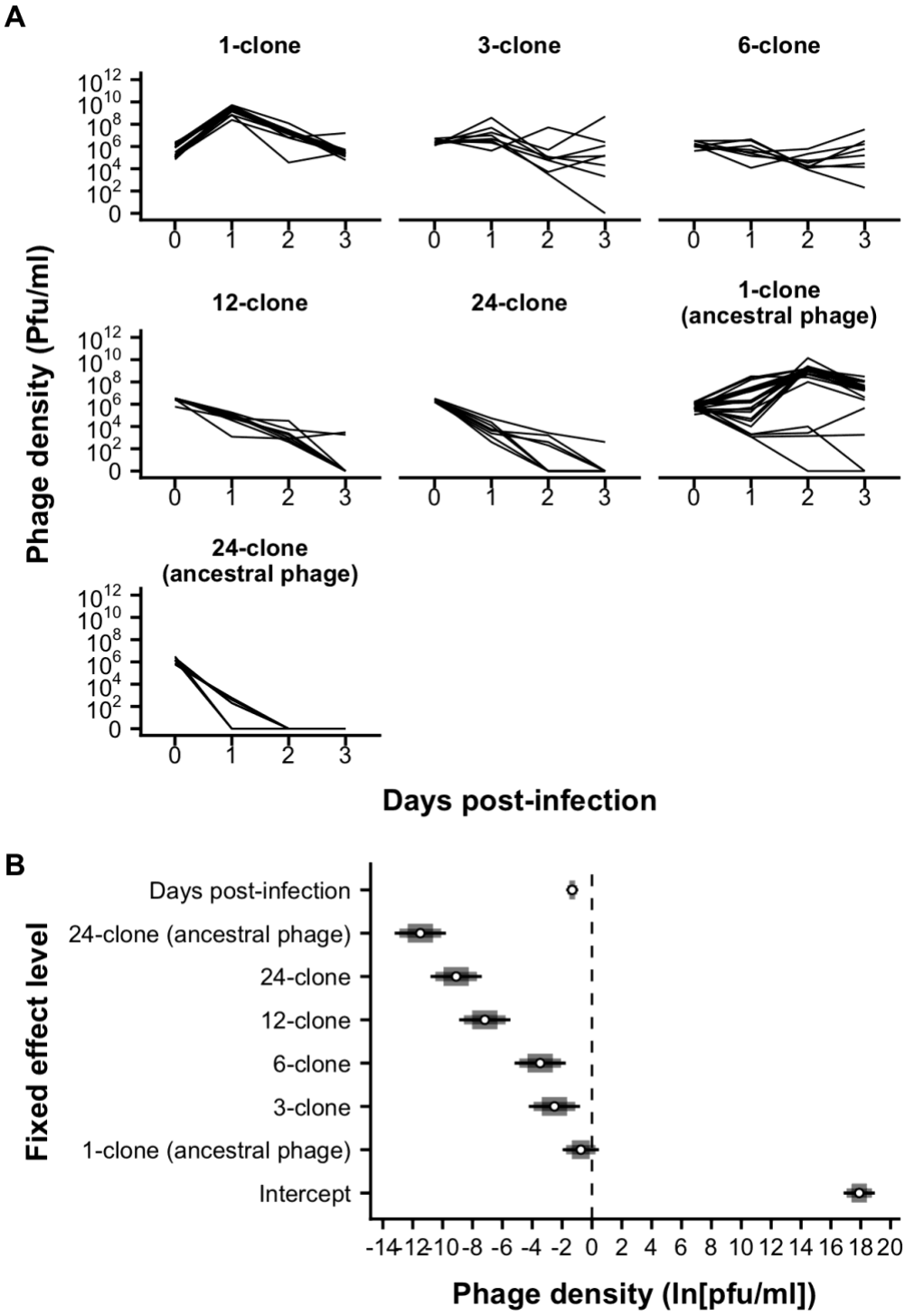


Figure 3.2. Increasing CRISPR diversity limited phage population size

(**A**) Population dynamics of phage at different levels of CRISPR diversity in the host population (panels). Black lines show the phage density expressed as plaque-forming units (pfu) ml^{-1} in individual replicates at each day post-infection. The limit of phage

(Figure 3.2 caption contd.)

detection is 200 pfu ml^{-1}. (**B**) Coefficients from a GLMM of the natural log of phage density (ln[pfu ml^{-1}]), with CRISPR allele diversity treatment and days post-infection (dpi) as fixed effects. The intercept is the mean phage titre in the 1-clone treatment at 0 dpi. The β differences from the intercept (shown as a dotted line) are given for the remaining levels of both fixed effects. Means are shown as white points with 67, 89% and 95% credibility intervals given in decreasing width.

susceptible fraction of the host population in the 24-clone treatment infected with escape phage.

Next, we hypothesised that if phage population size across our CRISPR diversity treatments was related to phage evolution, the proportion of phage that evolved an expanded infectivity range would correlate with CRISPR diversity. Given that our experiment was designed such that the proportion of the host population resistant to the escape phage increased with CRISPR diversity, there would likely have been strong selection to acquire mutations in other protospacers and PAMs to infect other hosts in the population. The evolution of range expansion did depend on diversity (Likelihood Ratio = 6.60, p = 0.01), and was most likely in the 6-clone treatment, particularly at 3 days post-infection (dpi) (**Fig. 3.3**). This result suggests that intermediate host diversity in the 6-clone treatment increased the likelihood that novel escape phage with an expanded host range would evolve (Benmayor *et al.*, 2009; Chabas *et al.*, 2018). The fact that these novel phage could infect more hosts likely explains why phage population size was more consistent in the 6-clone compared to the 3-clone treatment (**Fig. 3.2A**), and there was no significant difference in phage population size between these treatments (**Fig. 3.2B**). Together, these data suggest that the increased likelihood of evolutionary emergence at intermediate host diversity balances against the potential limitation of reproduction caused by the dilution effect. These novel phage variants could then maintain phage population size.

Additionally, the presence of phage in some treatments was promoted by evolution of a host shift, where infectivity on the original host is lost but phage evolved to infect

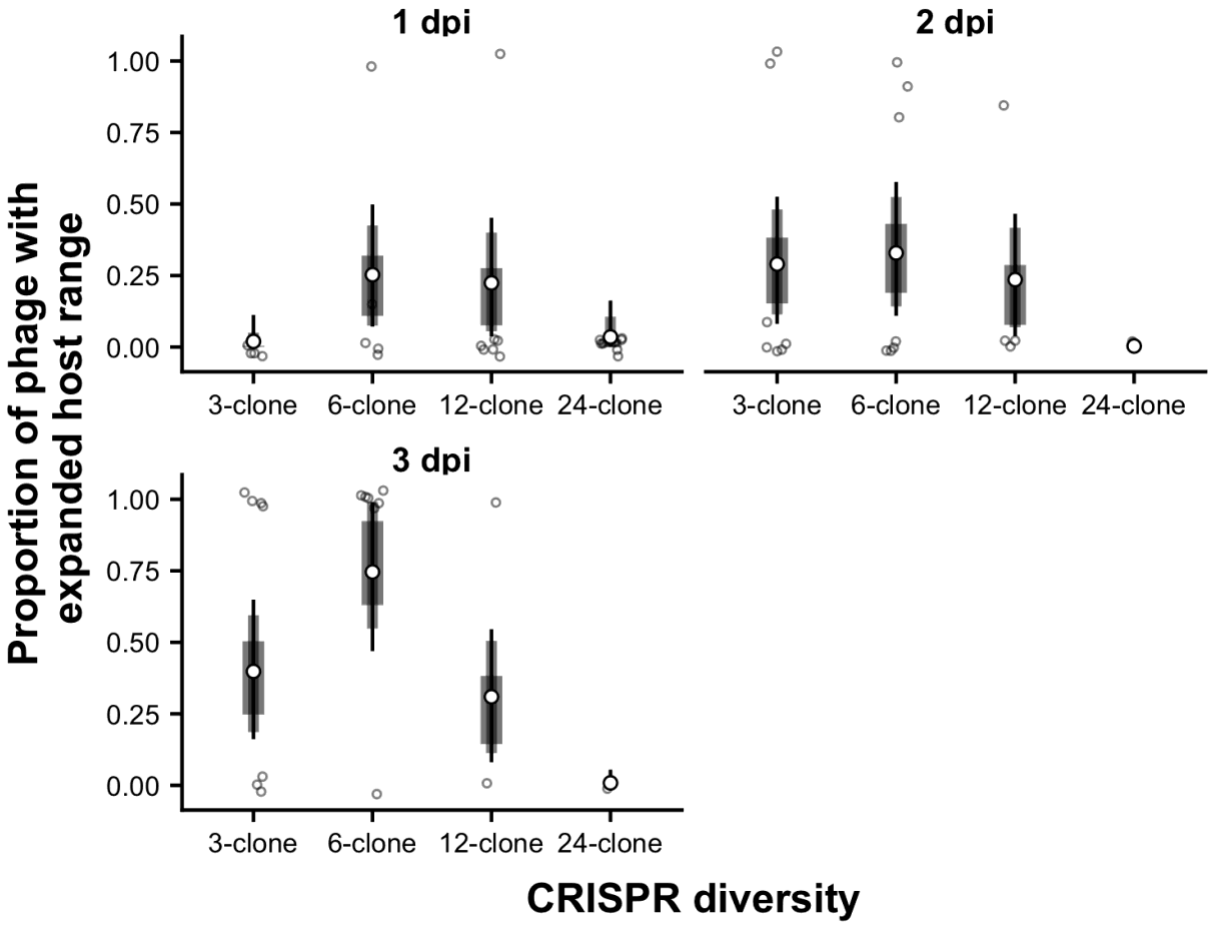


Figure 3.3 Phage evolutionary emergence was maximised at intermediate CRISPR diversity

Estimated proportion of phage that had evolved to infect a second CRISPR clone in addition to the original CRISPR clone they were pre-evolved to infect at each day post-infection in each polyclonal CRISPR diversity treatment. Because these estimates are derived from a Bayesian GLMM, they are posterior modes - not means – and are shown as white points. Similarly, 67, 89 and 95% credibility intervals are highest probability density intervals, and are given in decreasing width around mode estimates. Raw values from each replicate are shown as points.

a new host. This is a less likely event as it involves two mutations: back-mutation to the ancestral state at the original protospacer followed by mutation at the new protospacer, all while incurring the cost of loss of infectivity on the original host. We observed phage in two out of eight replicates of the 24-clone treatment had evolved a host shift. Sequence data confirmed that single nucleotide polymorphisms (SNPs) in the PAM that conferred infectivity to the original host had been lost, and novel SNPs or deletions had emerged in protospacers that conferred infectivity to a new host present in the population (**Tables A3.4 & A3.5**). Recall that phage titres in the 24-clone (ancestral phage) treatment were lower than the 24-clone treatment infected with escape phage (**Fig. 3.2B**). Together with the presence of phage that evolved by host shift, these data suggest that the establishment of a modest phage epidemic in this treatment may have facilitated the evolutionary emergence of novel phage.

Given that CRISPR diversity negatively affected phage population size, we expected that this would lead to enhanced fitness of all bacterial hosts with CRISPR immunity relative to hosts without CRISPR immunity (van Houte *et al.*, 2016b). This is because the fitness costs associated with CRISPR immunity are conditional upon phage exposure (Westra *et al.*, 2015). We therefore hypothesised that the selection rate (Lenski, 1991; Travisano & Lenski, 1996) of all CRISPR clones (both resistant and susceptible) compared to the Δ*pilA* strain (which does not use CRISPR immunity) would be positively related to CRISPR diversity. Despite some variation among replicates, we did find that CRISPR clones in all polyclonal host populations had higher selection rates compared to clonal populations during infection with escape phage (**Figs. 3.4A & B**). In 3, 3, and 4 replicates in the 12-, 24-, and 24-clone

(ancestral phage) treatments, respectively, the density of the Δ*pilA* strain dropped to zero (**Figure A3.1)**, suggesting that hosts with CRISPR immunity could occasionally outcompete those without. However, the selection rate of hosts with CRISPR immunity did not differ among the polyclonal treatments ($F_{4,113}$ = 1.73, p = 0.15; **Fig. 3.4C**). Selection rate also did not change notably over time (regression of selection rate over dpi: β [95% CI] = 0.03 [-0.04, 0.11], $t(2)_{118}$ = 0.94, p = 0.35) even though phage densities decreased with time.

The absence of a detectable relationship between the level of CRISPR diversity and the selection rate of hosts with CRISPR immunity relative to those without might be because escape phage dynamics were dependent on proportion of sensitive bacteria in the host population. We hypothesised that the selection rate of susceptible CRISPR clones relative to resistant CRISPR clones would be positively related to the level of CRISPR diversity, because the relative frequency of sensitive hosts declines as CRISPR diversity increases. Indeed, in some replicates in the 3- and 6-clone treatments the selection rate of the susceptible CRISPR clone was <-6, indicating that they were driven extinct by phage (**Fig. 3.5A**). Further, the selection rate of susceptible CRISPR clones in these treatments was on average less than zero (mean selection rate [95% CI]; 3-clone: -1.53 [-2.61, -0.45]; 6-clone: -0.99 [-2.07, -0.09]), indicating they tended to be outcompeted by resistant CRISPR clones. Turning to the 12- and 24-clone treatments, the selection rate of susceptible clones was statistically similar to zero (12-clone: 0.12 [-0.96, 1.21]; 24-clone: 0.15 [-0.93, 1.24]; **Fig. 3.5B & C**), and higher than the 3- and 6-clone treatments (**Fig. 3.5B & C**). Together, these data show that the selection rate of susceptible CRISPR clones increased with CRISPR diversity. There was also an increase in selection rate of

susceptible CRISPR clones over time among polyclonal treatments (regression of selection rate over dpi: β [95% CI] = 0.41 [0.03, 0.80], $t(2)_{118} = 2.17$, $p = 0.03$; **Fig. 3.5A & B**), which parallels the decline in phage population sizes over time in the same treatments (regression of ln pfu ml^{-1} over dpi: β [95% CI] = -2.58 [-3.23, -1.84], $t(2)_{114} = -7.64$, $p = 7.27 \times 10^{-12}$). These data suggest that susceptible hosts are protected by population-level CRISPR diversity, and that this protection is mediated by a reduced ability of phage to replicate on increasingly dilute susceptible hosts (Payne *et al.*, 2018)(**Figure A3.1**).

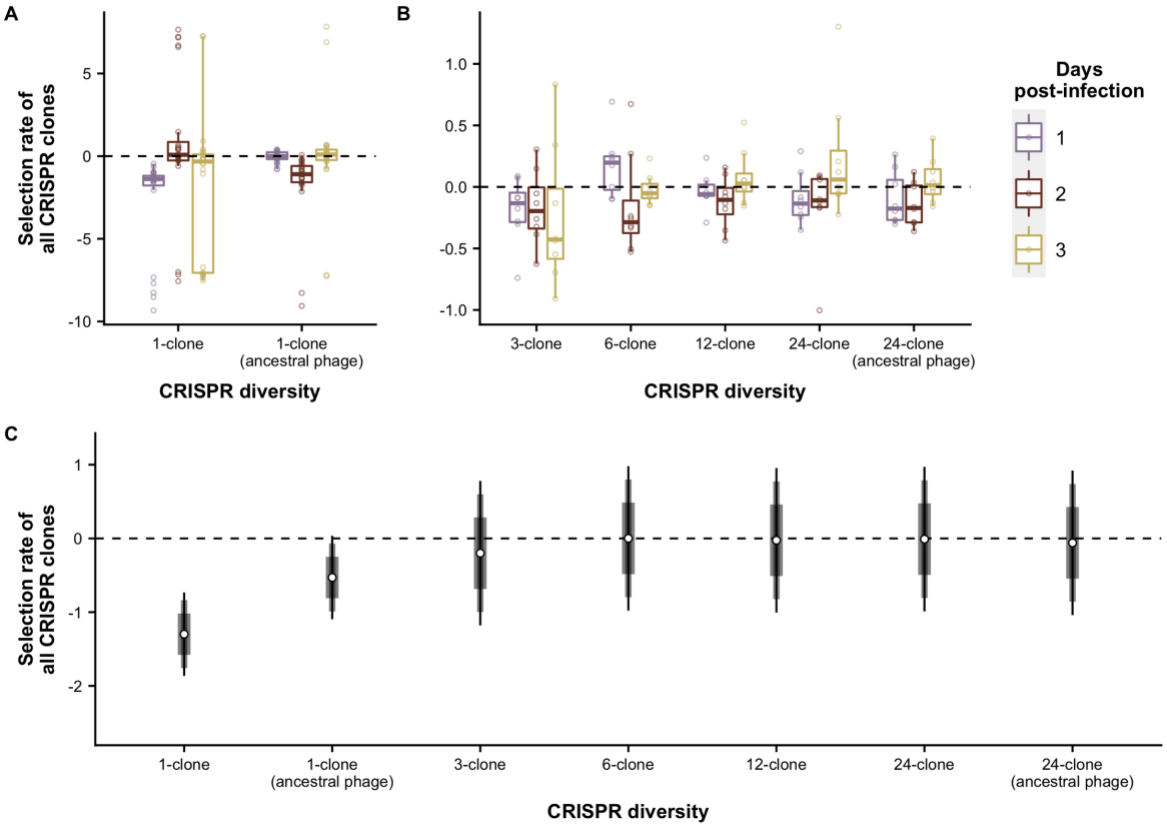


Figure 3.4 Selection rate of bacterial hosts with CRISPR immunity did not increase with CRISPR diversity in polyclonal treatments

Selection rate of hosts with CRISPR immunity relative to a strain without CRISPR immunity in (**A**) the 1-clone CRISPR diversity treatments with infective and ancestral phage and (**B**) the remaining CRISPR diversity treatments. Days post-infection is indicated by colour. Selection rate is the natural log of the relative change in density of bacteria with CRISPR immunity (both susceptible and resistant to phage) against a strain without. The dotted line at zero indicates no difference in density change i.e. both are equally fit. Boxplots show the median, 25th and 75th percentile, and the interquartile range. Raw values from each replicate are shown as points. (**C**) Predicted mean selection rate is calculated from a GLMM that statistically controls for the effect of time. Means are shown as white points with 67, 89 and 95% credibility intervals given in decreasing width.

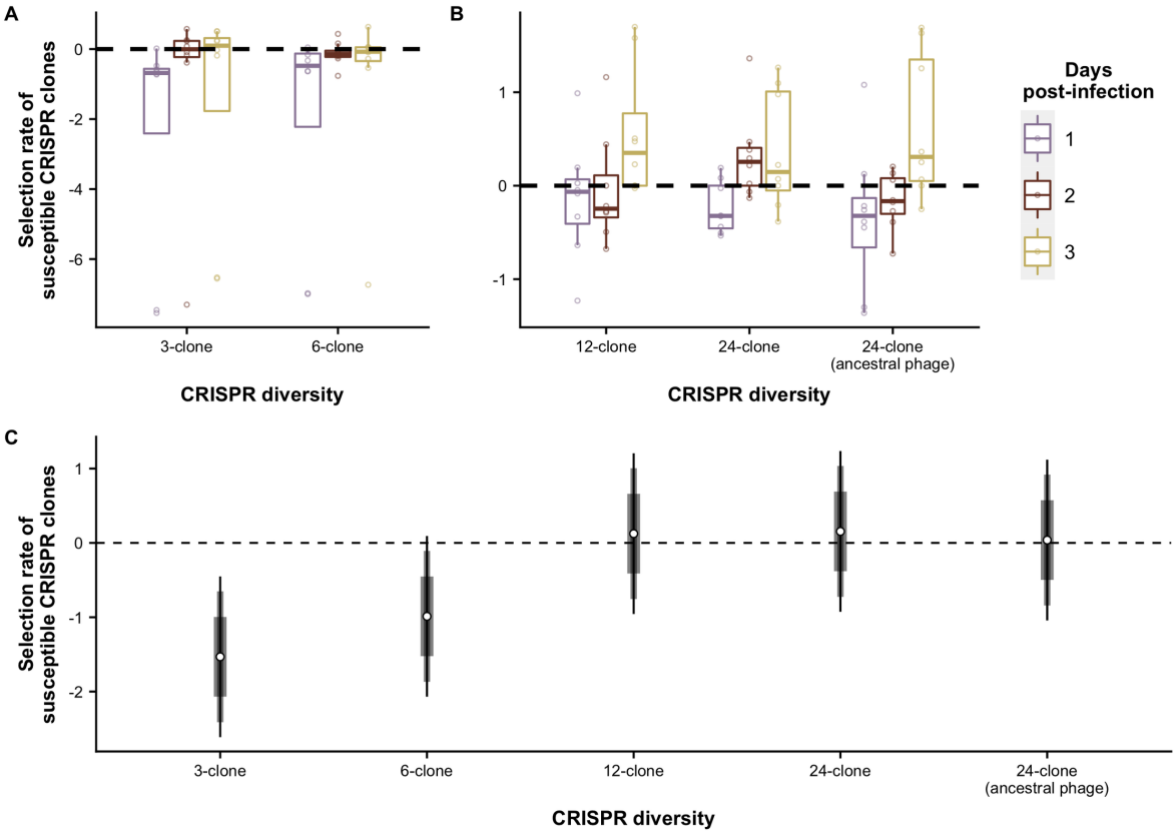


Figure 3.5 The selection rate of susceptible CRISPR clones increased with CRISPR diversity and over time

Selection rate of the susceptible CRISPR clones relative to the resistant CRISPR clones in (**A**) the 3- and 6-clone CRISPR diversity treatments and (**B**) the remaining CRISPR diversity treatments. Days post-infection is indicated by colour. Selection rate is the natural log of the relative change in density of CRISPR clones that were susceptible to phage against CRISPR clones that were resistant to phage. The dotted line at zero indicates no difference in density change i.e. both are equally fit. Boxplots show the median, 25th and 75th percentile, and the interquartile range. Raw values from each replicate are shown as points. The 1-clone treatments, both with infective and ancestral phage, have been excluded because all bacterial clones with CRISPR-based resistance in this treatment were susceptible, and the data is therefore present in Fig. 4A. (**C**) Predicted mean selection rate is calculated from a

(Figure 3.5 caption contd.)

GLMM that statistically controls for the effect of time. Means are shown as white points with 67, 89 and 95% credibility intervals given in decreasing width.

Discussion

Previous studies have shown that CRISPR diversity can limit the evolution of phage to overcome host resistance (van Houte *et al.*, 2016b; Chabas *et al.*, 2018). In those studies, bacterial populations were infected with ancestral phage that had not been previously exposed to resistant hosts. Here, we examined the consequences of CRISPR diversity once a phage has already evolved to overcome one of the CRISPR resistance alleles in the population. This tractable system enabled us to closely monitor how the level of CRISPR diversity influenced the population and evolutionary dynamics of the phage, as well as the evolutionary dynamics of the host.

These analyses show that phage population sizes were larger and persisted longer in populations with lower CRISPR diversity, which is consistent with previous studies (van Houte *et al.*, 2016b). Our results indicate that phage population size is increasingly limited by host diversity, because increasing host diversity reduces the proportion of susceptible individuals in the host population. This dilution effect of host resistance diversity is an important factor in explaining why genetically diverse host populations often have reduced pathogen loads (Keesing *et al.*, 2006; Ostfeld & Keesing, 2012; Civitello *et al.*, 2015). The finding that resistant hosts can protect the susceptible fraction of the population is also consistent with the epidemiological concept of “herd immunity” (Fine *et al.*, 2011). While herd immunity has been extensively studied in eukaryotes, recent work has suggested it may be important in understanding bacteria-phage dynamics (Payne *et al.*, 2018), a suggestion that our results provide further experimental support. This result also qualitatively matches

theoretical predictions of how host diversity should affect pathogen spread in a system with matching allele genetic architecture (Lively, 2010).

In theory, this kind of ecological effect can in turn shape the evolutionary dynamics of the bacteria-phage interaction, since smaller phage population sizes will decrease the mutation supply and hence the evolutionary potential of the phage (Antia *et al.*, 2003; Dennehy *et al.*, 2006; Yates *et al.*, 2006). We did observe that the evolution of phage host range expansion peaked at intermediate levels of CRISPR diversity. This is likely because increasing CRISPR diversity dilutes susceptible hosts, which results in smaller phage population size and hence less genetic variation on which selection can act. Simultaneously, increasing host diversity can increase selection for escape mutations. These two effects in dynamic tension can maximise evolutionary emergence at intermediate host diversity (Chabas *et al.*, 2018). The kind of interdependence between ecological and evolutionary dynamics of phage suggested by this study provides important context for other factors that can affect – and are affected by - evolutionary emergence. For example, the ability of a phage to escape a particular CRISPR spacer can influence host population structure by determining colony heterogeneity (Pyenson & Marraffini, 2020). Another example is when CRISPR immunity is incomplete, because phage with anti-CRISPR (*Acr*) genes can vary in the degree to which they inhibit the CRISPR immune response (Landsberger *et al.*, 2018). Differences in the degree of CRISPR immunity because of *Acr* diversity have been shown to affect the evolutionary emergence of escape phage (Chevallereau *et al.*, 2020). Understanding of how these factors interact with CRISPR allele diversity to affect the emergence and spread of escape phage will be important in future studies of CRISPR ecology and evolution.

In this study, we focussed on host populations where different CRISPR resistance genotypes were at equal starting frequencies, but natural communities are often composed of a few very common and many rare variants (Pachepsky *et al.*, 2001; McGill *et al.*, 2007). This likely matters for the observed dynamics, since the proportion of susceptible hosts has a large impact on the probability of evolutionary emergence of pathogens (Chabas *et al.*, 2018). Also, we focussed our analysis on the simple case where a diverse host population is infected by a clonal pathogen population. In nature, pathogen populations will frequently be genetically diverse as well (Hudson *et al.*, 2006; Telfer *et al.*, 2010), and increased levels of pathogen diversity may affect the benefits of host diversity (Ganz & Ebert, 2010). Indeed, previous studies of CRISPR-phage interactions suggest that infection by two different phages can increase bacteria-phage coexistence compared to infections with a single phage (Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015). Further, infection by diverse phage can select against CRISPR immunity in favour of surface-based resistance at the level of the host population, but also select for individual CRISPR clones with a more diverse array of spacers (Broniewski *et al.*, 2020). The empirical system used in this study offers both a unique ability to link genotypes and phenotypes, tight experimental control over the infectivity matrix of the host-phage interaction, and a clear link with theoretical predictions. These features will make it an ideal system for more detailed studies to understand how the composition of the host population and the relative diversity levels of the phage and host shape coevolutionary interactions.

Materials & Methods

Bacterial strains and phage

Evolution experiments were carried out using wild-type *Pseudomonas aeruginosa* UCBPP-PA14 (which has two CRISPR loci, CRISPR1 and CRISPR2), UCBPP-PA14 Δ*pilA* (this strain lacks the pilus, which is the phage DMS3 receptor, and therefore displays surface-based instead of CRISPR-based resistance) and phage DMS3vir (Zegans *et al.*, 2009). We used *P. aeruginosa* UCBPP-PA14 *csy3::lacZ* (Cady *et al.*, 2012), which carries an inactive CRISPR-Cas system, for phage amplification, and for top lawns in phage spot and plaque assays. *P. aeruginosa* PA14 *csy3::lacZ, Escherichia coli* DH5α (NEB), *E. coli* CC118 λpir (NEB), and *E. coli* MFD*pir* (Ferrieres *et al.*, 2010) were used for molecular cloning.

Co-culture experiment

To control the levels of CRISPR diversity in our evolution experiments, we established a library of 24 *P. aeruginosa* PA14 clones each carrying a single spacer in CRISPR2 (bacteriophage-insensitive mutants; BIMs). We also independently evolved 24 phage mutants that could infect each one BIM (escape phage).

We designed 5 treatments in which we manipulated the level of CRISPR spacer diversity, based on the BIM library: monocultures (1-clone), or polycultures consisting of 3, 6, 12 and 24 clones. In order to monitor the population dynamics and relative fitness of individual bacterial clones within the polyculture treatments over

the course of the experiment, we transformed 8 BIMs to carry a *lacZ* reporter gene. The *LacZ* gene encodes the β-galactosidase enzyme that hydrolyses X-gal, resulting in the production of a blue pigment. For each of the polyclonal treatments, a single BIM carrying the *lacZ* reporter gene was included. 8 BIMs were chosen for transformation so that a single clone could be monitored in each of the 3-clone mixtures (BIMs 1, 4, 7, 10, 13, 16, 19, and 22; see **Table A3.1**)

From fresh overnight cultures of each BIM, we made mixes of equal proportion of each clone corresponding to the diversity treatments. To monitor the population dynamics and competitive performance of all bacterial clones with an active CRISPR system, we also added PA14 Δ*pilA* (which is fully resistant to phage DMS3vir via surface-based resistance and has a distinct "smooth" colony morphology) to each mix in equal proportion to the CRISPR-carrying fraction of the population. We then inoculated 6ml of M9 minimal media (supplemented with 0.2% glucose; M9m) with bacteria from the mixed overnight cultures at a dilution of 1:100. Approximately $1x10^6$ pfu ml^{-1} of the escape phage targeting the labelled BIM were then added to each vial. We also established 1- and 24-clone treatments with ancestral phage as controls. Polyclonal treatments consisted of 8 biological replicates (*N*=8) to ensure that both BIMs and phage were equally represented across treatments, while the 1-clone treatments consisted of 24 biological replicates (*N*=24). Glass vials were incubated at 37˚C while shaking at 180rpm. At 1, 2, and 3 days-post infection (dpi), the sampling of the phage and bacterial culture was repeated as described below. Cultures were transferred 1:100 to fresh media after sampling had been carried out. The experiment was terminated at 3 dpi.

Sampling proceeded as follows. Each day 180µl of culture was taken from each vial and phage was extracted using chloroform. Phage titres were determined by serially diluting extracted phage in 1x M9 salts, and then spotting 5µl of each dilution on a top lawn of *P. aeruginosa* PA14 *csy3::lacZ*, which was then incubated at 37˚C for 24hrs. Phage titres were calculated from this assay. The detection limit of phage spot assays is 10^2 pfu ml^{-1}. To monitor bacterial densities, culture samples were serially diluted in 1x M9 salts, and then plated on LB agar + 40µg/ml X-gal + 0.1mM IPTG, and incubated for 48hrs at 37˚C. The density of SM, CRISPR and the labelled BIM was then calculated. SM were differentiated from CRISPR clones by their "smooth" colony morphology, and the labelled BIM was identified by the blue:white screen.

We assessed the competitive performance of the all BIMs relative to the Δ*pilA* strain, and the labelled BIM relative to non-labelled BIMs by calculating selection rate (*r*) each day (*t*) from 1-3 (*n*) dpi (r_n = (ln [density strain A at t_n/density strain A at t_{n-1}] – ln[density strain B at t_n/density strain B at t_{n-1}])/day) (Lenski, 1991; Travisano & Lenski, 1996), which expresses competitive performance as the natural log of the relative change in density of one competitor against another. We decided to calculate selection rate instead of relative fitness (*W*) as used in Westra *et al.* (2015), because *W* is very sensitive to starting densities and hence becomes increasingly skewed at later timepoints. Applied to our data, it consequently produced impossibly high values.

Phage evolution

We examined phage evolution during the experiment by sampling 12 individual plaques from each replicate that had detectable levels of phage from 1 to 3 dpi, which were amplified on PA14 *csy3::lacZ* overnight in LB, at 37˚C and shaking at 180rpm. Phage were extracted using chloroform, and then diluted 1000-fold. Samples of each phage were then applied on lawns of each of the 24 BIMs and PA14 *csy3::lacZ*. A successful infection was indicated by a clear lysis zone on the top lawn. Phage were classified according to whether they had expanded their infectivity range (could infect the original susceptible clone and a new clone in the BIM library). Of the phages that had undergone a host shift (lost infectivity to the original clone and could only infect a new clone), we confirmed their expanded infectivity range by sequencing the old and new protospacers on the evolved phage genome (SourceBioscience, UK). We also sequenced the relevant protospacers of the pre-evolved escape phage from the BIM-phage library and ancestral DMS3vir. Primers are given in Table S3.

Statistical analyses

All statistical analyses were carried out in R v3.5.3 (R Core Team, 2018). Generalised linear mixed models (GLMMs) were used throughout, and replicate was treated as a random effect in all models. Days post-infection (dpi) was treated as a continuous variable and CRISPR allele diversity as a categorical variable. Model selection followed a nested approach, where full versus reduced models were

compared using information criteria (Burnham & Anderson, 2003, 2004), as well as focussed comparisons of observed and predicted values. The overall statistical significance of the single and interaction effects of fixed effects (that is, treatment and dpi) was assessed using ANOVA, which gives an *F*- and associated *p*-value. Simple Bayesian GLMMs from the MCMCglmm (Hadfield, 2010) package were used to analyse phage evolution to overcome model conversion issues encountered with this dataset. These models used a probit transformation (the inverse standard normal distribution of the probability) and a flat prior. The overall statistical significance of the single and additive effects of fixed effects in these models was assessed using Likelihood Ratio (LR) tests and their associated *p*-values using the VCVglmm package (Brown, 2019). Where possible, exact *p*-values are given, but R is unable to give exact values when $p < 2.2 \times 10^{-16}$. When phage titre was considered as the response variable, data was log-transformed to improve model fit after adding 1 to every value. Credibility intervals around model coefficients and predictions were calculated to the 95%, 89% and 67% level to give the reader a clearer indication of effect size. All R code used for analyses and figures is available at this manuscript's GitHub - github.com/JackCommon/Common_etal_2020.

Acknowledgements

We thank Meaghan Castledine for help constructing the pBAM1(Gm)_lacZ plasmid.

Chapter 4: CRISPR-Cas immunity leads to a coevolutionary arms race between *Streptococcus thermophilus* and lytic phage

Publication details

The contents of this chapter have been published as the following:

> Common J, Morley D, Westra ER & van Houte S (2019) "CRISPR-Cas immunity leads to a coevolutionary arms race between *Streptococcus thermophilus* and lytic phage" *Philosophical Transactions of the Royal Society B* 374 DOI: 10.1098/rstb.2018.0098

The main text, figures, tables, and supplemental information are those which appear in publication. They have been formatted according to submission requirements and to provide consistency throughout the thesis. Details specific to publication, such as funding details and data accessibility, have been excluded. Author contributions can be found under Statement of contributions as co-author.

Abstract

CRISPR-Cas is an adaptive prokaryotic immune system that prevents phage infection. By incorporating phage-derived “spacer” sequences into CRISPR loci on the host genome, future infections from the same phage genotype can be recognised and the phage genome cleaved. However, phage can escape CRISPR degradation by mutating the sequence targeted by the spacer, allowing them to re-infect previously CRISPR-immune hosts, and theoretically leading to coevolution. Previous studies have shown that phage can persist over long periods in populations of *Streptococcus thermophilus* that can acquire CRISPR-Cas immunity, but it has remained less clear whether this coexistence was due to coevolution, and if so, what type of coevolutionary dynamics were involved. In this study, we performed highly replicated serial transfer experiments over 30 days with *S. thermophilus* and a lytic phage. Using a combination of phenotypic and genotypic data, we show that CRISPR-mediated resistance and phage infectivity coevolved over time following an arms race dynamic, and that asymmetry between phage infectivity and host resistance within this system eventually causes phage extinction. This work provides further insight in the way CRISPR-Cas systems shape the population and coevolutionary dynamics of bacteria-phage interactions.

Introduction

Clustered Regularly Interspaced Short Palindromic Repeats and their associated *cas* genes (CRISPR-Cas) form an adaptive immune system that is found in approximately 50% of all bacteria and 90% of archaea (Grissa *et al.*, 2007a). CRISPR-Cas confers immunity to phage infection by incorporating phage-derived sequences into CRISPR loci on the host genome. These loci consist of repeating sequences ("repeats") that are interspaced by sequences ("spacers") derived from phage and other mobile genetic elements of typically around 30nt in length. RNA transcripts of CRISPR loci are processed and form a ribonucleoprotein complex with Cas proteins that can recognise and cleave complementary nucleic acid sequences, preventing future infections by the same phage genotype. CRISPR-Cas systems are highly diverse, and are currently ordered into two classes, six types and 33 subtypes based on their *cas* gene composition, gene synteny and CRISPR repeat sequences, with clear differences in the molecular mechanisms of different variants (Koonin *et al.*, 2017).

In some natural environments, bacteria with CRISPR-Cas systems appear to coevolve with phage over long time periods (Andersson & Banfield, 2008; Laanto *et al.*, 2017). However, studying the dynamics of these coevolutionary interactions under controlled laboratory conditions has been limited by the availability of adequate model systems. Specifically, while many bacteria encode CRISPR-Cas immune systems, under laboratory conditions the vast majority do not evolve CRISPR-based immunity upon phage or plasmid infection, or do so at such low frequencies that they are detectable only with deep-sequencing approaches. Such

low-frequency CRISPR evolution is unlikely to significantly contribute to the reciprocal selection between the bacteria and the phage. Currently, only two bacterial species have been found to naturally evolve (almost) exclusively CRISPR-based immunity under laboratory conditions: *Streptococcus thermophilus* strains DGCC7710 and LMD-9 (Deveau *et al.*, 2008; Horvath *et al.*, 2008; Horvath *et al.*, 2009), and *Pseudomonas aeruginosa* strain UCBPP-PA14 (Westra *et al.*, 2015).

Early studies with *S. thermophilus* demonstrated that phage can overcome CRISPR immunity by evolving point mutations in the sequence targeted by the spacer (the "protospacer"), or in the protospacer-adjacent motif (PAM) (Deveau *et al.*, 2008), a conserved sequence immediately adjacent to the protospacer that is used by the bacteria to discriminate between self (i.e. CRISPR loci) and non-self (i.e. phage) DNA (Mojica *et al.*, 2009; Semenova *et al.*, 2011; Westra *et al.*, 2013). This observation suggested a possible scenario for coevolution in free-running systems, where bacteria acquire spacers over time and phage escape via point mutations in the corresponding protospacers or PAMs (Childs *et al.*, 2012; Weinberger *et al.*, 2012; Iranzo *et al.*, 2013). Consistent with this idea, it was reported that *S. thermophilus* can coexist with phage over many generations, and that for each treatment the single experimental population displayed large fluctuations in its spacer repertoire and an increase in the frequency of point mutations in phage genomes over time (Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015; Sun *et al.*, 2016). However, a more recent study suggested that coevolution is unable to explain long-term coexistence of *S. thermophilus* and its phage, suggesting instead that this may be driven by back mutation of bacteria with CRISPR immunity to sensitive phenotypes (Weissman *et al.*, 2018a), which would provide a continuous supply of

sensitive hosts for phage to amplify in. Such loss of CRISPR immunity due to mutation has also been observed at high frequencies in *Staphylococcus epidermidis* (W. Jiang *et al.*, 2013), and reversion to sensitive phenotypes more generally may be an important mechanism for bacteria-phage coexistence (Chaudhry *et al.*, 2018).

Given the lack of clarity surrounding the role, if any, and the type of CRISPR-phage coevolution for bacteria-phage coexistence, we performed highly replicated, long-term (30-day) serial transfer experiments with *S. thermophilus* and its lytic phage 2972. Our phenotypic assays demonstrate that bacteria and phage coevolved in these experiments during at least the first 9 days (approximately 70 generations). We next examined the type of coevolutionary dynamics during this period, with a clear distinction between fluctuating selection dynamics (FSD), where rare host and pathogen genotypes are favoured through negative frequency-dependent selection, and arms race dynamics (ARD), where host resistance and phage infectivity increase over time (Vale & Little, 2010). We found that CRISPR-mediated immunity and phage infectivity increase over time. Further, our genotypic data show that patterns of resistance and infectivity were explained by bacteria acquiring novel spacers against the phage, and the phage evolving mutations in the regions targeted by the spacers. Collectively, the data show that coevolution is likely an important factor in the coexistence of bacteria and phage in this empirical system, and that this coevolution is characterized by an arms race dynamic.

Experimental methods

Strains used in the study

We used the lactic acid bacterium *Streptococcus thermophilus* DGCC7710 wild-type (WT) and its lytic *pac*-type phage 2972 (GenBank: NC_007019.1) (Lévesque *et al.*, 2005) as a model system. DGCC7710 has four CRISPR-Cas systems, two of which (CRISPR1 and CRISPR3) are active during infection with phage 2972 and both are classified as Type II-A (Horvath *et al.*, 2008; Carte *et al.*, 2014). CRISPR1 has 32 spacers and CRISPR3 has 12 spacers (Grissa *et al.*, 2007b; Horvath *et al.*, 2008), none of which are perfectly complementary to a PAM-flanked sequence in the phage 2972 genome.

Phage 2972 amplification

An overnight culture of *S. thermophilus* was transferred 1:10 into fresh LM17 medium (M17 broth supplemented with 0.5% α-Lactose) containing 10mM $CaCl_2$ and incubated shaking at 180 revolutions per minute (rpm) at 42˚C. When the culture reached log phase (OD_{600} ~ 0.25) approximately 10^6 plaque forming units (pfu) of phage 2972 were added and the culture was incubated under the same conditions for two hours, at which point cells had fully lysed. Lysates were centrifuged and filtered through a 0.22 µm filter, and the resulting phage stocks were stored at 4˚C.

Long-term co-culture experiment

Prior to commencing the experiment, *S. thermophilus* was acclimatised in LM17 medium at 42˚C and 180 rpm for 2 days, with a 1:100 transfer into fresh LM17 after 24 hours (approximately 10^6 cfu). To start the co-culture experiment, overnight cultures of the bacteria were transferred 1:100 into 6ml LM17 media supplemented with 10mM $CaCl_2$ in glass vials. They were then infected with either 10^9, 10^8, 10^7 or 10^6 pfu of phage 2972, with 12 independent replicate experiments per treatment, followed by incubation at 42˚C while shaking at 180 rpm. Replicates were transferred 1:100 into fresh LM17 + 10mM $CaCl_2$ every 24 hours and phage titres and bacterial densities were measured every 24 hours for a period of 30 days, or until no phage was detected for four consecutive days. Bacterial densities were determined through plating and colony counts, while phage densities were measured by plaque assays. These were performed by mixing phage dilutions with WT bacteria in soft agar overlays (LM17 + 10mM $CaCl_2$ and 0.5% agar), poured onto hard agar (LM17 + 10mM $CaCl_2$ and 1.5% agar).

Phage survival

Phage survival and mean time to extinction over the course of the experiment were analysed using a Cox proportional hazards model from the survival package (Therneau & Lumley, 2015).

Measuring the evolution of infectivity and resistance

To measure whether host resistance and phage infectivity evolved during their co-culture, we isolated phage clones and bacterial clones from the treatment where

bacteria were infected with 10^8 pfu phage. To give sufficient power in this analysis, we focussed on eight replicate experiments from this treatment where phage persisted for at least nine days, a period which we estimated to be sufficiently long for significant coevolution to take place. Phage extracted from 1, 4 and 9 days post-infection (dpi) were subjected to plaque assays as described above. Three time points were chosen as a minimal requirement for the downstream time shift analysis to monitor whether coevolution took place and to determine the coevolutionary dynamics (see below). For each replicate and time point, twelve plaques were randomly picked and amplified in 96 well plates containing LM17 + 10mM $CaCl_2$ in which WT bacteria were inoculated 1:100 from a fresh overnight culture. Bacteria extracted from the same time points were diluted and plated overnight, and twelve colonies from each replicate were picked at random and used to make soft agar overlays on LM17 agar lawns. To examine the evolution of phage infectivity for each of the eight replicates, the 36 phage clones that were isolated (12 phage clones x 3 time points) were stamped using a 96-pin replicator onto 36 bacterial lawns corresponding to the bacterial clones isolated from the same replicate (i.e. 12 bacterial clones x 3 time points). Phage were classified as being infective against a particular bacterial clone if a clear lysis zone was visible on the lawn after incubation at 42˚C for 24 hours. If no lysis zone was visible, the host was classified as resistant.

Evolution of infectivity and resistance

Using the data from the experiments described above, we measured the evolution of phage infectivity as the proportion of bacterial clones that phage from each time point from the same replicate experiment could infect (i.e. how phage infectivity

range changed over time). In a similar way, we measured the evolution of host resistance as the proportion of all phage genotypes from the same replicate experiment that could be resisted by bacteria from each time point (i.e. how host resistance range changed over time). Infectivity or resistance was analysed in a Generalised Linear Model (GLM) with genotype as a fixed effect and a binomial family with a logit link function. Mean infectivity or resistance was then analysed for each time point in a Generalised Linear Mixed Model (GLMM) using the lme4 package (Bates, 2015), with time point as a fixed effect and replicate as a random effect. Model coefficients and confidence intervals were transformed from logits to probabilities prior to presentation.

Time-shift experiment

Because the susceptibility and resistance of bacterial clones to phage from past, present or future time points was determined (**Table 4.1**), our phenotypic assay also served as a time-shift experiment (Gaba & Ebert, 2009). Time-shift experiments involve challenging samples of host or pathogen populations from a particular time point against samples of pathogen or host populations from contemporary, past and future host or pathogen populations. Time-shift experiments are a powerful tool to characterise underlying coevolutionary processes, and have been used in several host-pathogen systems (Koskella & Lively, 2007; Gandon *et al.*, 2008), including bacteria-phage (Buckling & Rainey, 2002a; Hall *et al.*, 2011b; Koskella, 2014). Our phage infectivity and host resistance data was analysed as a time-shift experiment by first scoring each pairwise challenge as ‘past’, ‘present’, or ‘future’, with reference to the phage’s background compared to the host. Infectivity was then analysed in a

GLMM with phage background as a fixed effect and replicate, host timepoint, and phage timepoint as random effects. Models had a binomial family with a logit link function.

To test for the relative importance of arms race dynamics (ARD) versus fluctuating selection dynamics (FSD), we estimated the strength of the Genotype x Environment (GxE) effect on infectivity and resistance following Hall *et al.* (2011b). Under a simple arms race, all hosts should be more susceptible to phage from their future compared to their past or present, independent of genotype. Environment (E) therefore refers to the time point from which phage originate in pairwise challenges. By contrast, under fluctuating selection different host genotypes will vary in their susceptibility to hosts from their past, present or future. Measuring which proportion of the variation in susceptibility across phage environments can be explained by the interaction between the environment and host genotype (G) can therefore be used to measure the relative contribution of FSD. Increasing values of this proportion (GxE / E) relate to increasing differences in susceptibility among host genotypes. We estimated this by calculating the ratio of the mean square (MS) of an Environment-only model to the MS of a GxE model for each replicate at each time point. These ratios were then analysed in a GLMM with time point as a fixed effect and replicate as a random effect, with a normal family and square root link function.

Statistical methods

For all experiments, statistical analyses were carried out in R v3.5.0 (R Core Team, 2018), and graphics were generated using r-base and the ggplot2 package

(Wickham, 2009). Model selection followed a nested design, and the final models in all analyses were selected based on the reduction of heteroskedacity, χ^2 tests, and AIC comparisons (Akaike, 1973; Burnham & Anderson, 2003, 2004). Where appropriate, tests were Bonferroni adjusted using the multicomp package (Hothorn *et al.*, 2008).

Spacer sequence analysis

For all bacterial clones that were isolated from the eight replicate experiments where bacteria had been infected with 10^8 pfu of 2972, expansion of the CRISPR1 and CRISPR3 arrays was analysed using colony PCR to determine whether spacer acquisition had taken place (12 clones x 3 time points x 8 replicates = 288 clones in total). The following primers were used: CRISPR1 5'-tgctgagacaacctagtctctc-3' and 5'-taaacagagcctccctatcc-3'; CRISPR3 5'-ctgagattaatagtgcgattacg-3' and 5'-gctggatattcgtataacatgtc-3'. Clones that had acquired new spacers were further analysed by Sanger sequencing of the amplicon (Source Bioscience, UK), followed by mapping of the spacers against the phage 2972 genome (Accession: NC 007019.1) using BLAST followed by manual verification with Geneious v9.1.8 (Kearse *et al.*, 2012). Spacer diversity was calculated as the pairwise difference (PWD) among nucleotides between spacer sequences. The effect of spacer number and diversity on infectivity was analysed in a GLMM, with either number or diversity as fixed effects and replicate as a random effect. Models had a binomial family and logit link function.

Phage sequence analysis

To understand whether phages could escape CRISPR immunity through target site (protospacer) mutation, we selected phage clones (out of the 288 total phage clones isolated, see above) and sequenced the protospacer(s) that would match the spacer(s) present among the 12 isolated bacterial clones in a replicate. We then analysed the protospacers and their associated protospacer-adjacent motifs (PAMs) for SNPs that could explain the ability of the phage to overcome CRISPR immunity. For phage clone selection we used the individual infectivity matrices generated from our phenotypic assays (**Table A4.1**), and we only sequenced phage clones from infection matrices that showed phage infectivity on hosts that had acquired spacers i.e. excluded from analysis were phages from those matrices where none of the 12 host clones of a replicate had acquired spacers or where none of the 12 phage clones of a replicate showed infectivity. Sequenced phage clones were taken from 1, 4 and 9 dpi. At least two individual phage clones were selected from each matrix that was analysed, based on their ability to infect CRISPR-immune hosts. When an infection matrix showed a high degree of variation between phage clones, more than two phage clones were analysed so that most of the variation in infectivity would be covered (e.g. in the matrix of replicate 7, T9 phage:T9 host, phage clones 1, 2, 5, 6 , 8, and 9 were selected [**Table A4.1**]). Where possible, one phage clone that did not show infectivity to any of the 12 bacterial clones from each matrix was taken along to serve as a control for protospacer sequencing, alongside an ancestral phage as control. Primers for protospacer sequencing were designed in Geneious by using the available spacer information (**Tables A4.2 & A4.3**). PCR amplicons of total of 51 phage clones were generated by performing PCR on the filtered phage stock, which

was followed by Sanger sequencing. To identify point mutations, sequences were first mapped to the 2972 genome using Geneious. Single nucleotide polymorphisms (SNPs) in either the seed sequence or PAM were identified (Deveau *et al.*, 2008; Horvath *et al.*, 2008), and SNP locations were then compared against the protospacer sequence targeted by the CRISPR array of each clone that phage had been challenged against (**Table A4.4**). Phage with SNP(s) in the seed sequence or PAM of the targeted protospacers were scored as 'predicted infective'. We found that approximately 70% (241/348) of predicted infectious phage were measured as successfully infecting a host (**Table 4.3**). The remaining predicted infections that were not measured as successful may be attributable to partial CRISPR resistance or other resistance mechanisms such as surface modification. We also expect some degree of experimental error in our assay given that detection of lysis zones is a relatively crude method of discerning infectivity/resistance. Using data from the phenotypic assay, we then analysed the effect of mutation on infectivity. The effect of escape by point mutation on infectivity was modelled in a GLMM as the proportion of infections associated with phage that had a SNP in the protospacer seed sequence or PAM. We analysed the effect of the evolution of the number of SNPs in all targeted sequences matching the host's CRISPR array by first subtracting the number of targeted protospacers that had evolved from the total number of spacers in each host. This gives the number of targeted sequences that had not evolved. The proportion of infections was then modelled against these values. All models included replicate as a random effect, and used a binomial family with a logit link function.

Results

We set out to first examine the generality of the previously reported population dynamics following infection of *S. thermophilus* DGCC7710 with a single phage 2972. We therefore infected 12 replicate experimental populations of *S. thermophilus* DGCC7710 with either 10^6, 10^7, 10^8 or 10^9 plaque forming units (pfu) of phage 2972 (12 independent replicates per treatment; 48 populations in total), and monitored the bacterial and phage population densities on a daily basis for 30 days. For the first three days following infection, phage titres remained fairly constant in most replicates between 10^6 -10^8 pfu ml^{-1}, with the exception of the highest phage treatment (10^9 pfu) where phage and bacteria went extinct in 11 out of 12 replicates (**Fig. 4.1**). Lower phage titres were correlated with higher host densities (z = -0.31, 95% CI = -0.42, -0.19, $p < 0.0001$). With the exception of the 10^9 treatment, this relationship between phage and host titres was the same among treatments ($F_{2,551} = 2.24$, $p = 0.11$). At 16 days post-infection (dpi), the phage had gone extinct in 44 / 48 replicates, and phage persisted for the entire 30-day duration of the experiment in two replicates, one each in the 10^7 and 10^8 pfu treatments. For the treatments where bacteria survived, the mean time until phage extinction in days was as follows: for the 10^9 pfu treatment: 2 ± 0.54 days; 10^8 treatment: 11.50 ± 1.77 days; 10^7 treatment: 11.50 ± 2.12 days; and 10^6 treatment: 7.67 ± 1.67 days (mean ± standard error).

Using these experimental lines, we first determined whether the coexisting bacteria and phage had evolved during their co-culture. Since the population dynamics associated with the 10^6-10^8 pfu infection regimes was virtually identical, we decided to limit our downstream analyses to the 10^8 treatment only. Further, to achieve

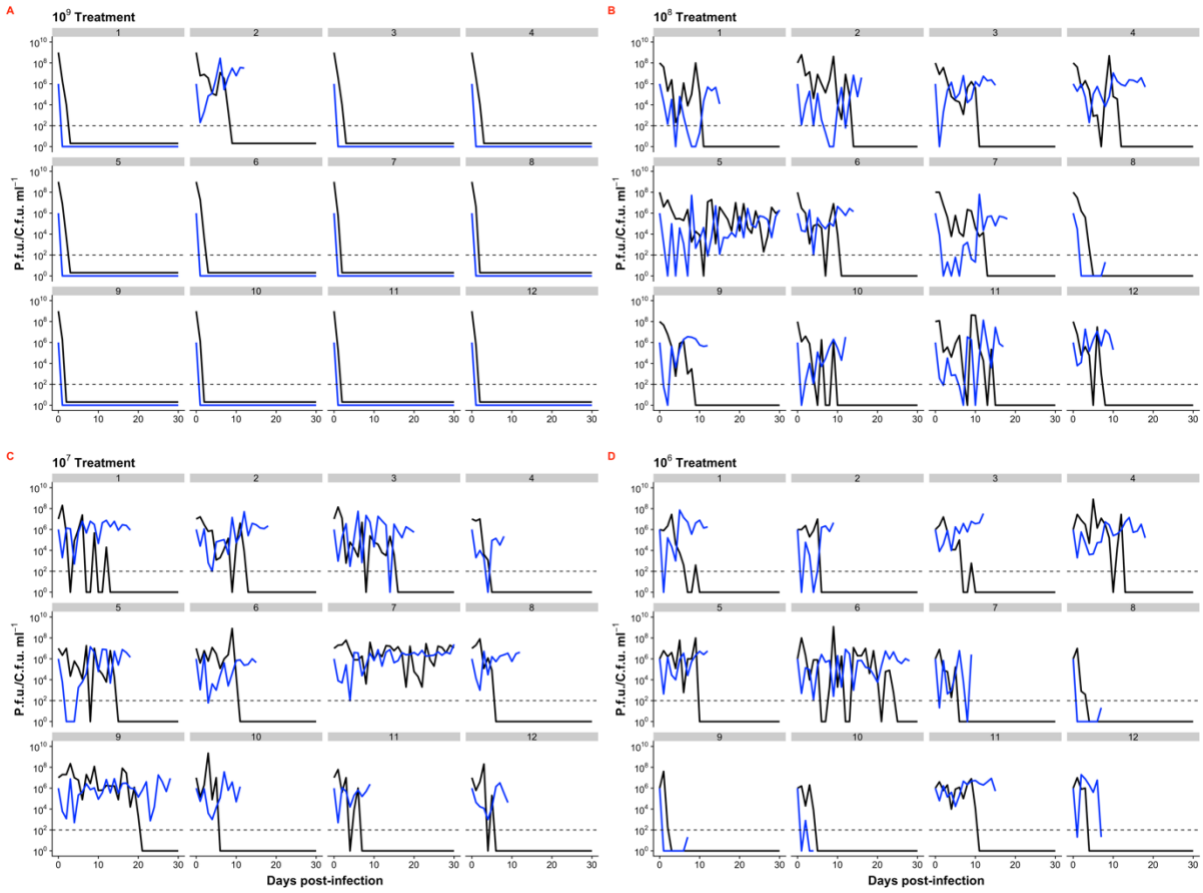


Figure 4.1 Phage and host population dynamics over time in each replicate

A-D) 10^9-10^6 pfu phage treatments, respectively, with replicate identity indicated above each sub-panel. Phage titres (plaque-forming units; pfu ml^{-1}) are shown in black and host densities (colony-forming units; cfu ml^{-1}) are shown in blue. The level of detection is 200 pfu ml^{-1} (dashed line).

sufficient power in our analyses and deduce coevolutionary dynamics, we selected replicates where bacteria and phage coexisted for at least nine days, resulting in a total of eight replicate populations that were examined in detail (**Fig. 4.1**). We then isolated 12 bacterial clones and 12 phage clones from each replicate at 1, 4 and 9 days post-infection (dpi). Using the 288 phage and 288 bacterial isolates, we first examined whether the phage and bacteria had evolved increased infectivity and resistance over time. This was done by measuring the resistance of each individual bacterial clone against all phage clones derived from the same replicate, and measuring the infectivity of each individual phage clone against all bacterial clones from the same replicate. This analysis revealed that mean phage infectivity (the proportion of all host genotypes that can be infected by a given phage genotype) increased significantly from 0.29 (CI = 0.08, 0.48) at 1 dpi to 0.57 (CI = 0.37, 0.74) at 4 dpi, but remained stable at 0.53 (CI = 0.33, 0.74) from 4 to 9 dpi. Mean host resistance (the proportion of all phage genotypes resisted by a given host genotype) increased significantly each time point, from 0.01 (CI = 0.00, 0.05) at 1 dpi to 0.67 (CI = 0.18, 0.96) at 4 dpi, and to 0.99 (CI = 0.96, 0.99) at 9 dpi (**Fig. 4.2**). Collectively, these data show that bacteria evolved to resist essentially all phage genotypes by 9 dpi, but phage did not evolve high levels of infectivity to match.

Having established that bacteria evolved increasing resistance and that phage evolved increasing infectivity over time, we next examined whether both species coevolved; and if so, what type of coevolutionary dynamics were associated with this system. To answer this question we performed a phenotypic time-shift experiment whereby bacteria were exposed to phage from their past, present and future (Gaba & Ebert, 2009; Koskella, 2014) which enabled us to measure infectivity and

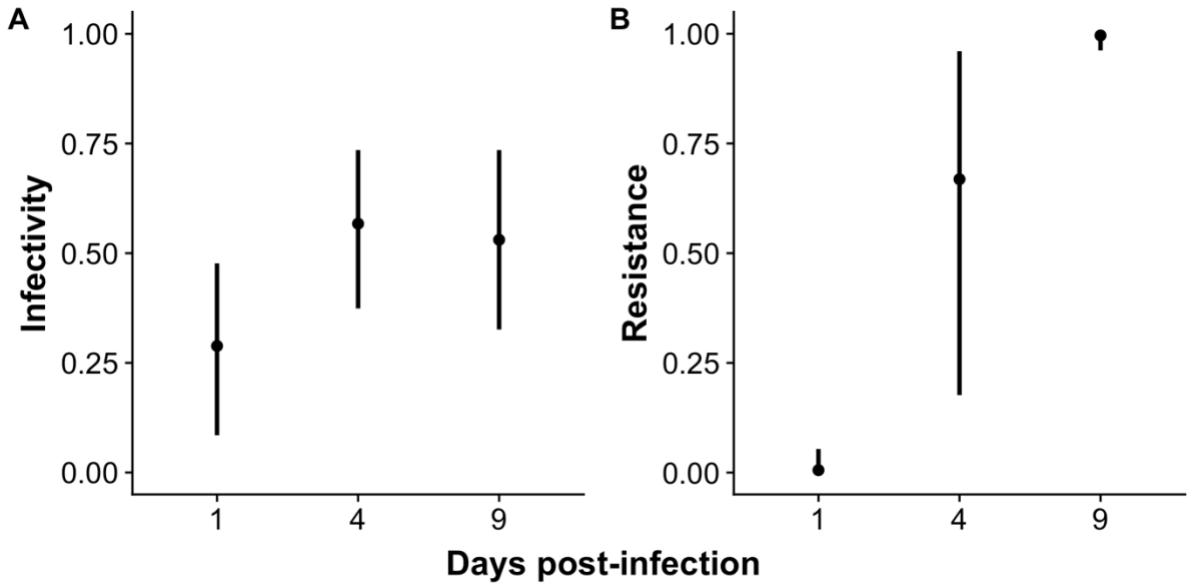


Figure 4.2 Evolution of infectivity and resistance over time

A) Phage infectivity over time, represented as the proportion of host genotypes from all time points that were infected by a given phage genotype in each replicate. **B)** Host resistance over time, represented as the proportion of phage genotypes from all time points that were resisted by a host genotype in each replicate. Means and 95% CIs are shown (*N* = 8).

		Phage		
		1	**4**	**9**
Host	**1**	Present	Future	Future
	4	Past	Present	Future
	9	Past	Past	Present

Table 4.1 Pairwise challenges between phage and hosts in the time shift assay

Numbers indicate the time points (days post-infection) analysed. Past, present or future refer to if hosts were contemporaneous or not with respect to the phage.

resistance patterns over time. Because individual genotypes may differ in their response to time-shift challenges, generalized linear mixed models (GLMMs) with replicate, host timepoint, and phage timepoint as random effects and phage background (**Tab. 4.1**) as a fixed effect were used to analyse time-shift data. This analysis showed that the original time-point of the phage with respect to the host had a significant effect on infectivity ($\chi^2_{(4,10044)} = 5.35$, $p < 0.0001$, $R^2 = 0.88$). Hosts were least susceptible to infection from past phage, more susceptible to contemporaneous phage, and most susceptible to phage from their future **(Fig. 4.3A)**. This pattern of increasing susceptibility from past to future phage generally held true when each pairwise combination of host and phage time point was considered (**Figure A4.1** & **Tab. 4.2**). Finally, host susceptibility to phage from the same time point declined consistently from 1 to 9 dpi (**Fig. 4.3B**). These data are consistent with an arms race dynamic (ARD) where hosts and pathogens escalate resistance or infectivity over time, but one in which host resistance eventually outpaces pathogen infectivity.

We formally tested for the relative importance of arms race versus fluctuating selection in our experiment by estimating the strength of the genotype X environment (GxE) effect on infectivity and resistance (Hall *et al.*, 2011b). Stronger GxE effects are consistent with stronger fluctuating selection (see *Experimental Methods*). This analysis showed that variation among genotypes was weak, consistent with a limited GxE effect (**Figure A4.2A & B**). The strength of the GxE effect did not change significantly with respect to time point for either phage infectivity ($\chi^2_{(2,24)} = 1.93$, $p = 0.38$, $R^2 = 0.13$) or host resistance ($\chi^2_{(2,24)} = 1.46$, $p = 0.48$, $R^2 = 0.11$). Collectively, these data demonstrate that *S. thermophilus* DGCC7710 and phage 2972 coevolved

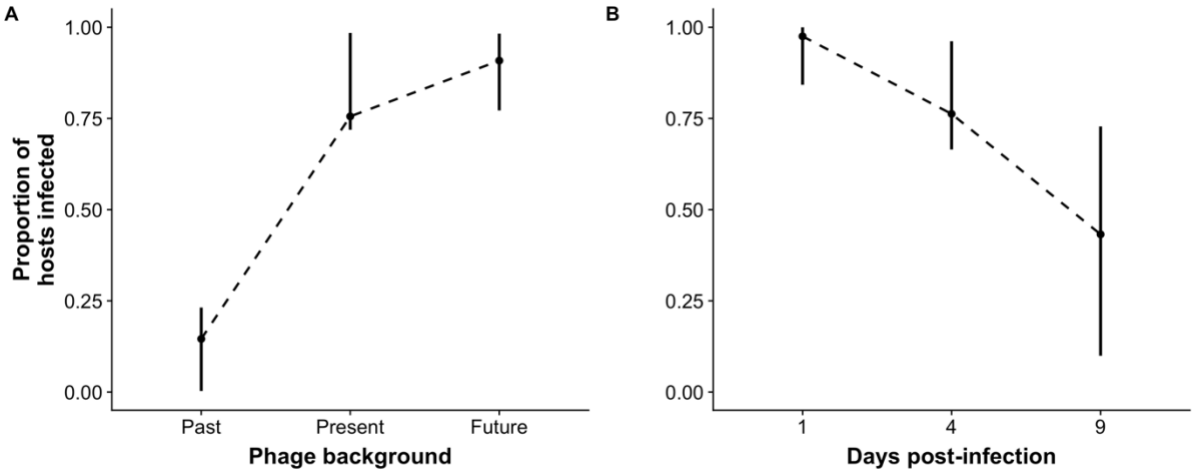


Figure 4.3 Results from time-shift experiment

A) Proportion of hosts infected when phage were from the host's past, present or future. **B)** Proportion of hosts infected by phage from the same time point (days post-infection). The dotted line is for illustrative purposes. Means are shown. 95% CIs represent the variation of the mean among replicates. (*N* = 100048).

Host origin (day)	Phage origin (day)	Mean Infectivity	Infectivity 95% CI
1	1	0.98	0.80---1.00
4	1	0.17	0.01---0.20
9	1	0	0.00—0.00
1	4	0.90	0.80---1.00
4	4	0.74	0.67---0.97
9	4	0.01	0.00---0.19
1	9	0.85	0.80---0.99
4	9	0.79	0.78---0.98
9	9	0.36	0.07---0.56

Table 4.2 Mean proportion and 95% confidence interval (CI) of hosts infected and phage resisted in pairwise challenges in the time-shift experiment, broken down by the day from which the host or phage originated. Values are rounded to two decimal places.

under these experimental conditions, and that the dynamics of their coevolution predominantly follows an arms race.

Based on previous studies showing that *S. thermophilus* typically acquires spacers in response to phage exposure (Deveau *et al.*, 2008; Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015), we predicted that this ARD was driven by reciprocal adaptation of the hosts' CRISPR array and the phage protospacers it targets. To test this, we first performed PCR analysis on the CRISPR1 and CRISPR3 loci of each bacterial clone to verify that the mechanistic basis of resistance was in fact due to the acquisition of novel CRISPR spacers. This revealed that the mean number of spacers per clone increased over time ($\chi^2_{(6,1140)} = 32.9$, $p < 0.0001$) **(Fig. 4.4A)**, and that all clones sequenced at 9 dpi had acquired at least one spacer (M = 0.55, CI = 0.45, 0.65) **(Fig. 4.4B)**. Comparison of these spacer sequences with the 2972 genome confirmed that they were acquired from the phage.

Using these sequencing data we determined the level of spacer diversity that naturally evolved within each replicate, since this is an important determinant of CRISPR-phage coevolution (Childs *et al.*, 2014; van Houte *et al.*, 2016b). Consistent with deep sequencing analyses of previous co-culture experiments (Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015), our data showed that spacer diversity, measured as the pairwise difference (PWD, with 0 indicating all spacers were shared among clones, and 1 indicating all spacers were unique) among spacer sequences, was generally low (grand mean = 0.25). Despite this, there was clear qualitative variation in spacer diversity between different replicates (**Fig. 4.4C**). Mean CRISPR genotype richness – the number of different CRISPR alleles we detected - was also

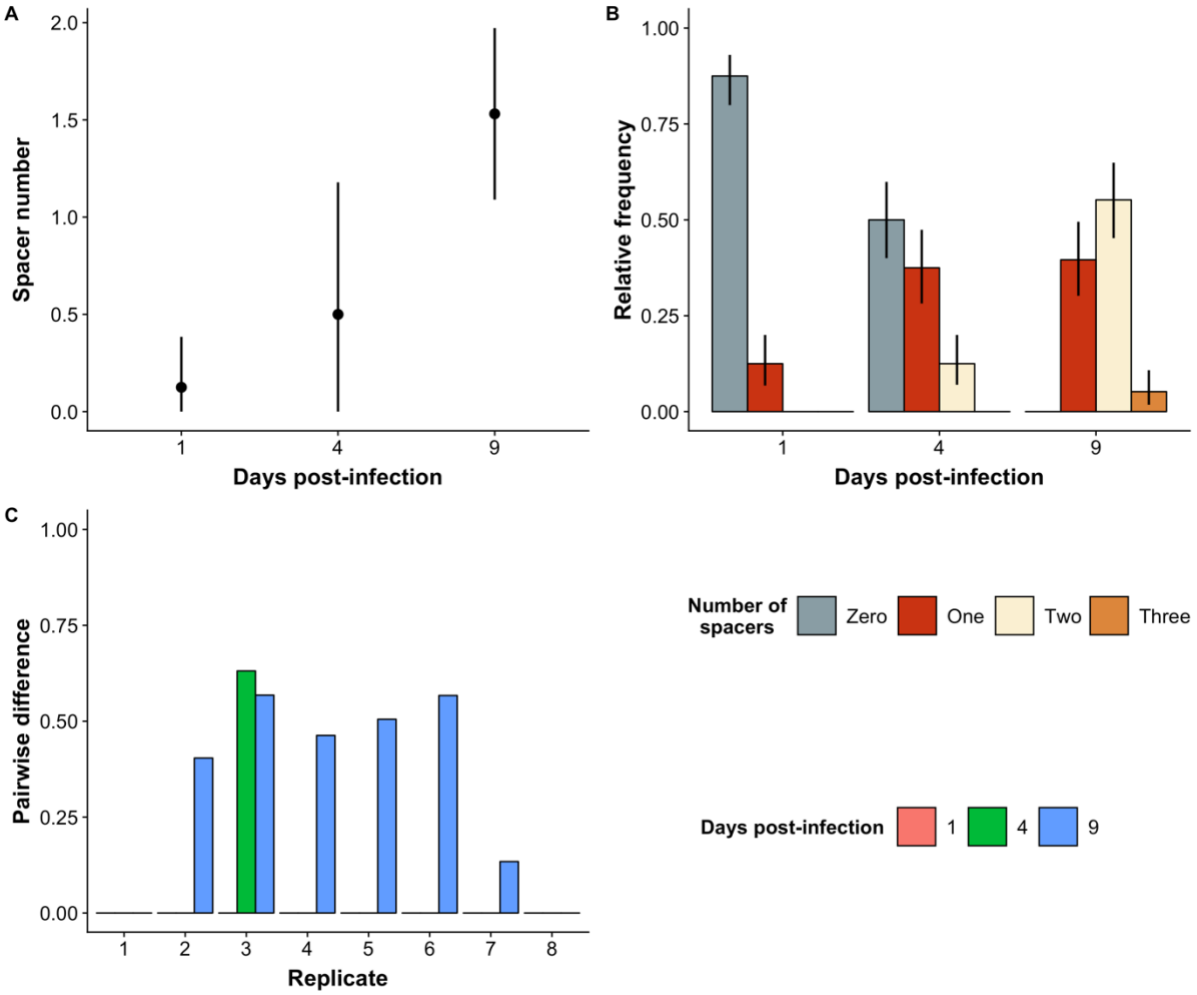


Figure 4.4 Spacers acquired during coexistence of *S. thermophilus* and phage 2972

A) Number of acquired spacers per clone at each day post-infection (dpi). Means and 95% CIs are shown (*N* = 8). **B)** Mean relative frequency of clones with different numbers of acquired spacers at each dpi. No clone with >3 spacers was detected. Means and 95% CIs are shown (*N* = 8). **C)** Spacer diversity in each replicate, measured as the pairwise differences among spacer sequences in each replicate (x-axis) at each time point (colours).

low, but increasing, across the sampled time points (1 dpi = 1, 4 dpi = 1.5, 9 dpi = 2.25). The diversity patterns become especially apparent when the spacers are mapped against the phage genome (**Fig. 4.5**) which shows that the spacer composition between time points can change completely, suggestive of selective sweeps of the population.

Consistent with previous theory and data (Childs *et al.*, 2014; van Houte *et al.*, 2016b) we found that host resistance increased with both the number of acquired spacers ($C = 1.91$, $z = 17.22$, $p < 0.0001$, $R^2 = 0.66$) (**Figure A4.3A**) and sequence diversity in terms of PWD ($C = 5.26$, $z = 0.27$, $p < 0.0001$, $R^2 = 0.44$) (**Figure A4.3B**). These data demonstrate that all clones that had acquired resistance had also acquired at least one novel spacer in either CRISPR1 or CRISPR3, suggesting that resistance is CRISPR-mediated. Further, Sanger sequencing of all CRISPR amplicons confirmed that all spacers that had been acquired indeed targeted the phage 2972 genome (**Table A4.2**). Spacers most frequently mapped to phage genes encoding hypothetical proteins compared to proteins with known functions, and the targeted genes tended to be at the distal end of the phage genome (**Table A4.2**). We next tested the hypothesis that the coevolutionary arms race we observed in these experiments was caused by reciprocal adaptation of the phage through the acquisition of point mutations in the target sequences. Such mutations have been observed in a previous co-culture experiment (Paez-Espino *et al.*, 2015), and provide a known mechanism for phage to overcome CRISPR resistance (Deveau *et al.*, 2008). To examine whether phage infectivity could be explained by the acquisition of point mutations, we first selected from the phenotypic assays a representative number of 56 different phage clones with different infectivity patterns

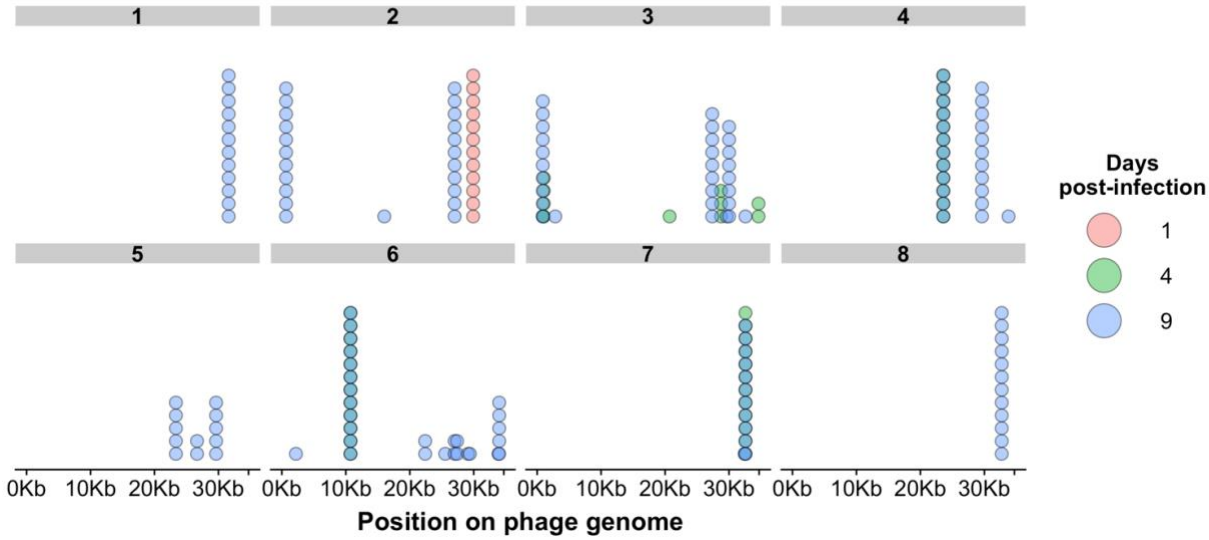


Figure 4.5 Locations of newly-acquired spacers on the phage 2972 genome

Histogram showing the location of acquired spacers in each replicate when mapped against the phage 2972 genome. Each dot represents a clone that had a spacer mapped to that region. Red, green and blue indicate 1, 4, and 9 days post-infection, respectively. Darker colours are the result of visual overlap between dots. Replicate identity is indicated above each sub-panel.

		Predicted	
		-	+
Measured	-	242	107
	+	106	241

Table 4.3 Contingency table of pairwise infections that were predicted to lead to phage escape based on protospacer sequence data

Single nucleotide polymorphisms (SNPs) in either the seed sequence or PAM were identified, and SNP locations were then compared against the protospacer sequence targeted by the CRISPR array of each clone that phage had been challenged against. These were then compared against the pairwise infections measured in the phenotypic assay. “+” indicates a successful infection, “-“ indicates no infection.

(i.e. covering both infective and non-infective phenotypes) across the three time points included in the phenotypic assay. We then PCR amplified their protospacer sequences based on the CRISPR spacer sequence data, followed by Sanger sequencing of the amplicons. This showed that 38 out of 51 selected phage clones had acquired at least one single nucleotide polymorphism (SNP) in the protospacer sequence or PAM (**Fig. 4.6A**)**;** the majority (33) of which were protospacer mutations (**Fig. 4.6B)**. Further, the majority (~70%) of phage that were predicted to be infectious based on sequence data were able to successfully infect hosts (**Tab. 4.3**), indicating that SNPs in the protospacer or PAM were generally sufficient to confer infectivity. The proportion of sequenced phage clones without SNPs in the protospacer or PAM from 4 dpi (10/22) was higher than phage clones from 9 dpi (4/26) (**Table A4.3**), which is consistent with the idea that CRISPR drives mutation of phage genomes in this empirical system (Paez-Espino *et al.*, 2015). Crucially, analysis of the infectivity patterns of sequenced phage showed that the mean number of infected hosts was significantly higher when phage had a SNP in the protospacer sequence or PAM compared to phage with no detectable mutations ($\chi^2_{(1,696)} = 32.22$, $p < 0.0001$, $R^2 = 0.31$) (**Fig. 4.6C**). Finally, phage that had evolved SNPs in all sequences that were targeted by the host's CRISPR array (0 targeted sequences) had significantly higher infectivity compared to phage that carried one or more unmutated target sequences ($\chi^2_{(1,696)} = 59.29$ $p < 0.0001$, $R^2 = 0.74$) (**Fig. 4.6D**). These data demonstrate that the acquisition of point mutations in the protospacer sequence in response to evolution of CRISPR immunity is the primary mechanism of phage reciprocal adaptation, driving the increase in phage infectivity during the observed arms race dynamic.

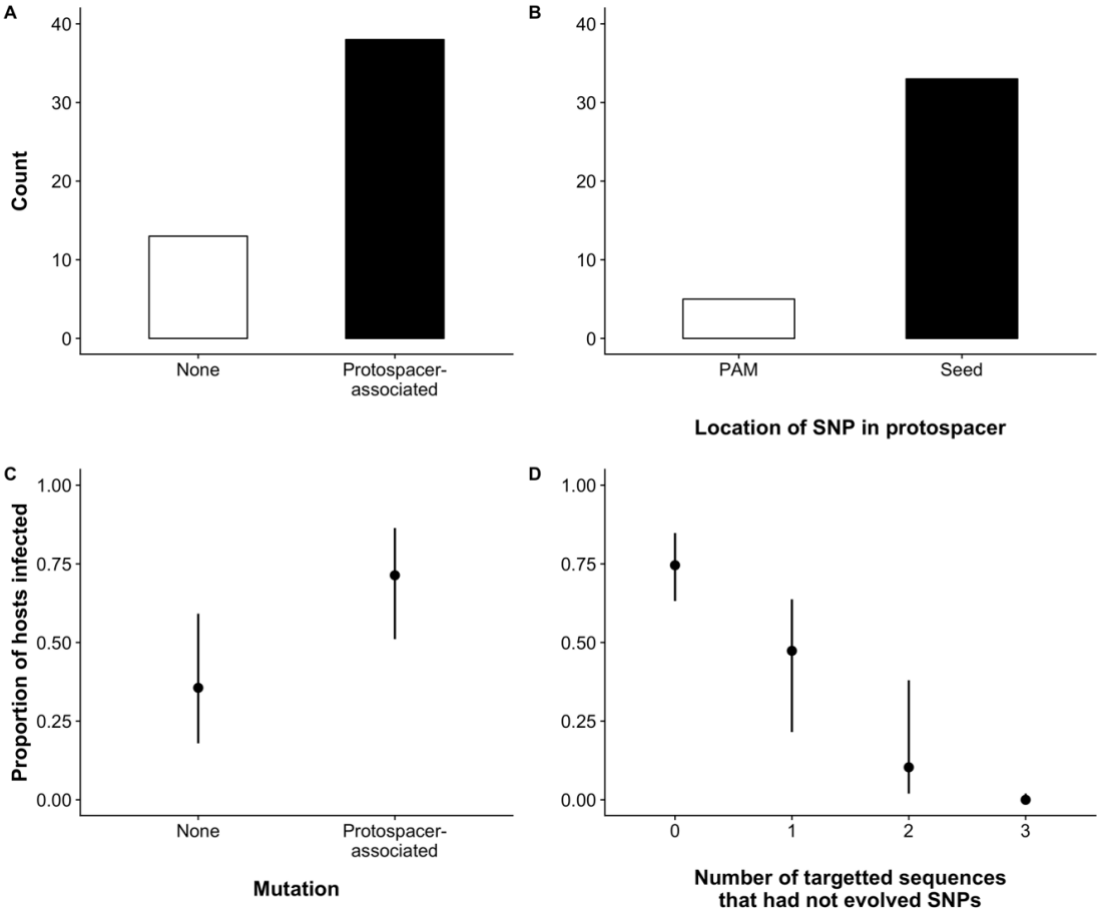


Figure 4.6 Protospacer sequence analysis and infectivity patterns

A) Histogram showing the number of sequenced phage (out of 56) that did not have a detectable mutation, had a 'random' mutation outside of the protospacer, or had a 'protospacer-associated' mutation either in the protospacer-adjacent motif (PAM) or the seed sequence. **B)** Number of sequenced phage with protospacer-associated mutations (out of 40) that had a single nucleotide polymorphism (SNP) in either the PAM or seed sequence. **C)** Mean proportion of hosts infected by phage that did or did not have a protospacer-associated SNP. The effect of random mutations is not included due to limited sample size. **D)** Mean proportion of hosts infected by phage that had a full or partial match to the host's CRISPR array in terms of the number of targeted protospacer sequences that had not evolved by point mutation. Means and 95% CIs are shown (N = 696).

Discussion

S. thermophilus DGCC7710 readily evolves CRISPR-based resistance in response to phage 2972 through spacer acquisition in two active CRISPR loci (CRISPR1 and CRISPR3). In return, phage can escape CRISPR immunity by evolving point mutations in the protospacer sequences targeted by CRISPR. This mechanism of host resistance and pathogen infectivity suggests a possible scenario for coevolution, where bacteria acquire spacers over time and phage accumulate escape mutations.

Consistent with earlier work on *S. thermophilus* and phage 2972 (Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015; Sun *et al.*, 2016; Weissman *et al.*, 2018a), we found that, with the exception of the highest initial phage concentration treatment, phage can coexist with bacteria over many generations despite the presence of CRISPR-based host immunity. Furthermore, our phenotypic data using bacteria and phages isolated during the first 9 dpi demonstrated that they coevolved following an arms race dynamic (ARD), with hosts and phage evolving increased resistance and infectivity over time, and hosts being more resistant to phage from the past compared to present and future time points. To the best of our knowledge, evolution of phage resistance is exclusively driven by CRISPR-Cas in this empirical system (i.e. surface modification has not been reported in *S. thermophilus* DGCC7710 in response to phage infections, and we have also never observed it in our experiments). Consistent with this, we found that the underlying mechanism of coexistence during this time span appears to be predominantly reciprocal adaptation of the hosts' CRISPR array and the phage protospacers it targets. Analysis of hosts'

CRISPR arrays show that they readily acquire phage-derived spacers, that hosts acquire more spacers over time, and that host resistance is strongly associated with both spacer acquisition and spacer number. In turn, phage evolved via point mutations in the targeted protospacers. Correlating this with our phenotypic data shows that such escape phage were on average more infective. Further, we find that phage had evolved SNPs in all target sequences matching the host's CRISPR array were most infective compared to those with an incomplete match.

It is notable that while hosts evolve resistance against essentially all phage, this is not matched by similarly broad phage infectivity range; phage at the last time point (9 dpi) could infect just over half of all hosts. The infectivity of contemporary phage also declines with time, suggesting that the evolution of host resistance "outpaces" that of phage infectivity. This asymmetry between host resistance and phage infectivity is consistent with the idea that bacterial hosts are 'ahead' in coevolutionary arms races (Buckling & Brockhurst, 2012). While asymmetrical arms races in other studied bacteria-phage systems are generally driven by a binary shift to a phage-resistant surface mutant (Lenski & Levin, 1985; Westra *et al.*, 2015), CRISPR-phage interactions suggest an alternative. Hosts can acquire multiple novel spacers with only a marginal cost (Vale *et al.*, 2015), but phage mutation is limited by mutation supply (Lenski & Levin, 1985; Levin *et al.*, 2013; Chabas *et al.*, 2018; Chabas *et al.*, 2019). In addition, full phage infectivity requires mutations in all the protospacers targeted by the host CRISPR array, which becomes increasingly difficult when individual hosts and populations acquire a greater number and diversity of spacer sequences over time (Levin *et al.*, 2013; van Houte *et al.*, 2016b). It is likely that this asymmetry leads to the repeatable phage extinctions we observed.

Interestingly, this and previous studies occasionally found quasi-stable long-term coexistence of bacteria and phage (Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015; Sun *et al.*, 2016; Weissman *et al.*, 2018a). Previous work suggests that this may be driven by back-mutation of resistant hosts towards sensitivity (Weissman *et al.*, 2018a). These two mechanisms for bacteria-phage coexistence may operate in parallel, and their relative importance remains to be investigated. The relative importance of coevolution for phage persistence contrasts with what is observed for *P. aeruginosa* and its phage DMS3vir, where phage are unable to coevolve with the host due to the high levels of spacer diversity that naturally evolve (van Houte *et al.*, 2016b). In this system a continuous supply of sensitive hosts can allow for bacteria-phage coexistence (Chabas *et al.*, 2016; van Houte *et al.*, 2016b; Westra *et al.*, 2017).

Our data clearly shows that host genotype diversity richness (i.e. the number of hosts with different CRISPR arrays) increases over time. This data, together with the rapid and repeatable phage extinction after day 9 in our experiment, indirectly supports the idea that *S. thermophilus* hosts receive a synergistic benefit from population-level CRISPR diversity in the context of phage infection. Host diversity is a key determinant of pathogen spread [reviewed in Keesing *et al.* (2010); Chabas *et al.* (2016)], and previous work on *P. aeruginosa* PA14 and *S. thermophilus* has shown that population-level CRISPR spacer diversity can limit phage persistence (van Houte *et al.*, 2016b). An understanding that population diversity may provide such a benefit clarifies other results from our data. Although phage infectivity was still possible even with a partial match to the host's CRISPR array, in the context of a

mixed, polyclonal host population, such reduced infectivity may limit phage reproduction and transmission, sufficient to cause the rapid extinctions we observed.

In at least some natural environments, bacteria that evolve CRISPR resistance and the phage they target can coexist (Andersson & Banfield, 2008). This may be due to CRISPR-phage coevolution, as recently observed for a fish pathogen and its phage (Laanto *et al.*, 2017), but long-term coexistence may also be explained by various other ecological and evolutionary factors that are absent from our simple laboratory environments (van Houte *et al.*, 2016a). For example, previous experiments suggest that longer periods of bacteria-phage coexistence are reached when experimental treatments contained multiple different phages (Paez-Espino *et al.*, 2013; Paez-Espino *et al.*, 2015). Further, phage under these conditions were found to escape not only by mutation, but also by recombination (Paez-Espino *et al.*, 2015). This is consistent with observations from other natural environments where phage recombinants were correlated with CRISPR activity (Andersson & Banfield, 2008). These examples highlight how biotic and abiotic complexities may be key in shaping the ecological and evolutionary dynamics of host-pathogen interactions, which we are only starting to understand in the context of CRISPR-phage interactions.

Acknowledgments

We thank Angus Buckling for useful discussion throughout the project.

Chapter 5: General Discussion

General remarks

"Diversity is one of our greatest strengths" – Bernie Sanders

In this thesis, I have sought to examine the role of diversity in bacteria-phage ecology and evolution. Broadly, I have presented novel insights into why host diversity can limit pathogen spread, the synergy between the processes involved, and some of the coevolutionary consequences. First, the density of susceptible hosts is a key determinant of pathogen success. Second, the close relationship between the density of susceptible hosts and population diversity influences if and how coevolution can maintain host-pathogen coexistence. Finally, I consider the coevolutionary dynamics that can emerge in the context of a diversity-generating mechanism.

Generation and maintenance of CRISPR diversity

I have shown that greater diversity leads to less phage adaptation to overcome immunity, which increases host fitness. However, different CRISPR-Cas immune systems generate different levels of spacer diversity. For example, while the CRISPR-Cas immune system of *P. aeruginosa* PA14 evolves high diversity in response to DMS3vir, leading to rapid phage extinction, I found that *S. thermophilus* coevolves with 2972 because of low initial CRIPSR diversity. Why not evolve higher levels of diversity?

One of the key features of CRISPR-Cas immunity is that different spacers do not differ in their host fitness effects (Childs *et al.*, 2014), and so diversity is unlikely to be generated by trade-offs among different immune alleles. CRISPR diversity in simulated and natural settings instead seems to be characterized by multiple dominant CRISPR strains that periodically sweep to fixation and create diversified bacteria-phage communities (Childs *et al.*, 2012; Iranzo *et al.*, 2013). Evidently then, different degrees of competition among CRISPR strains are unlikely to explain differences in the *de novo* generation of CRISPR diversity.

An alternative explanation is that differences in phage pressure alter the degree of CRISPR diversity. Indeed, parasites are often invoked to explain the maintenance of diversity in host resistance alleles (Brockhurst *et al.*, 2014), and “hot spots” of parasite pressure can select for sexual reproduction as a diversity-generating mechanism (Gomulkiewicz *et al.*, 2000; Thompson, 2005; King *et al.*, 2011). A role for phage pressure in the degree of CRISPR diversity seems likely, as higher force of infection could result in more sampling of the phage genome via spacer acquisition. A caveat here is that, beyond a certain threshold, a host population’s CRISPR system is unable to clear phage rapidly enough for coexistence to occur, as seen in the 10^9 phage treatment in Chapter 4. Below this threshold, however, further study would be required to compare levels of CRISPR diversity between host populations inoculated with different concentrations of phage.

Another possible explanation for differences in CRISPR diversity that CRISPR systems may have evolved different spacer acquisition rates as a result of prevailing

conditions in their evolutionary history. Recently, several studies have considered how factors such as temperature and mutation rates may explain the presence or absence of CRISPR systems (Drake, 2009; Weinberger *et al.*, 2012; Weissman *et al.*, 2018b). There has been comparatively little research attention given to how biotic or abiotic factors might affect key aspects of CRISPR systems, in particular those related to diversity, such as the rates of spacer acquisition and loss. For example, the different evolutionary histories of *P. aeruginosa* PA14 and *S. thermophilus* could explain why the CRISPR system of the former generates more diversity than that of the latter. Future work should investigate these mechanistic differences among CRISPR systems, as this would provide more clarity as to role of diversity in CRISPR-phage interactions in terms of coexistence, coevolution, and clinical effects.

Host-pathogen specificity networks

The BIM-phage library is an example of matching-allele network: each host genotype can be infected by one pathogen genotype, and *vice versa*. Previous work has examined how different coevolutionary processes can structure the pattern of infection networks (Poullain *et al.*, 2008; Beckett & Williams, 2013; Fortuna *et al.*, 2019). In Chapter 3, I experimentally demonstrate the reverse – how the pattern of an infection network influences the coevolutionary process. Given then that CRISPR-phage systems can be used to test the coevolutionary consequences of infection networks, further study could be undertaken to understand how this model system might be extended.

Diversity in host resistance is captured in several theoretical models by assumptions regarding the structure of the host-pathogen infection network (Agrawal & Lively, 2002). By utilising the molecular basis of CRISPR-phage interactions, assumptions regarding the infection network could be replicated *in vitro* to validate model predictions. The effectiveness of validating models with experiments on CRISPR-phage is in fact demonstrated by Chabas *et al.* (2018). Another potential use for experimental CRISPR-phage infection networks is to replicate those found in nature. Many natural host-pathogen systems exhibit specificity networks, for example in *Daphnia magna* (Luijckx *et al.*, 2013), and marine archaea and their phage (Flores *et al*., 2013). Future experimental work could construct networks observed in natural host-pathogen systems using a CRISPR-phage library. With this model system, it would be possible to how different ecological conditions affect the coevolutionary consequences of infection networks, as well as provide insight into the genotypic and phenotypic changes that take place.

Metapopulation dynamics

The instances of phage evolving host shift in the most diverse treatments in Chapter 3 were possibly related to a small initial escape phage epidemic establishing on susceptible hosts, which allowed host shift to occur. Although host shift only led to transient phage survival in our experiment, the effect of a susceptible host fraction in the context of a diverse, mostly-resistant population may have implications for more complex host-pathogen systems. Importantly, the results I have presented in this thesis are derived from experiments where bacteria and phage interact in a well-mixed and relatively homogenous liquid culture. Realistically, the majority of hosts

and pathogens are patchily distributed into subpopulations. A question which warrants further study therefore is: how does spatial heterogeneity interact with host diversity, and what are the coevolutionary consequences?

Metapopulation dynamics in host-pathogen systems have already been shown to influence local adaptation (Morgan *et al.*, 2005), coevolutionary rate (Brockhurst *et al.*, 2007), host-pathogen coexistence (Schrag & Mittler, 1996), and the evolution of alternative immune mechanisms (Chabas *et al.*, 2016). Potentially, host subpopulations with intermediate levels of diversity might promote the evolution of host shift or host range expansion. These coevolutionary "hot-spots" (Gomulkiewicz *et al.*, 2000; Thompson, 2005) could then conceivably act as sources of novel pathogen genotypes. Future work should therefore examine how the effects of host diversity and metapopulation dynamics interact, for example in terms of how host immigration can alter subpopulation diversity and the resulting epidemiology and coevolutionary impacts on the host-pathogen metapopulation.

Pathogen diversity also matters

Less well-understood is how pathogen diversity may interact with host diversity. Like their hosts, pathogens are often diverse (Haldane, 1949; Hudson *et al.*, 2006; Telfer *et al.*, 2010). A small number of studies have investigated how host and pathogen diversity might interact and influence epidemiology and host success. Experiments with the water flea *Daphnia magna* and its microsporidium parasite *Octosporea bayeri* found that parasite diversity did not have a notable effect on parasite prevalence when interacting with mono- or polyclonal hosts (Altermatt & Ebert,

2008). However, a later study in a similar system found that diverse parasites were more successful in both mono- and polyclonal hosts (Ganz & Ebert, 2010). In bacteria-phage systems, infection with clonal or diverse Φ2 does not affect the likelihood of resistance evolution in *P. fluorescens* (Gorter *et al.*, 2015). In the context of CRISPR-Cas, infection by clonal or diverse phage can alter the evolutionary dynamics of bacterial immunity at the population- and individual-levels (Broniewski *et al.*, 2020). A recent meta-analysis suggested that there is not a significant interaction between host and pathogen diversity (Ekroth *et al.*, 2019), however all but one study analysed to derive this result did not make pathogen diversity explicit nor included it as a controlled variable. The effect of relative host-pathogen diversity, if any, on pathogen success is consequently poorly-understood and ambiguous.

Two further questions regarding pathogen diversity arise from the work presented in this thesis. First, how do relative levels of host and pathogen diversity influence the population dynamic and coevolutionary trajectories of host-pathogen systems? One hypothesis is that the dilution effect mediated by host diversity still limits a diverse pathogen population, as higher pathogen diversity cannot overcome the problem of decreased contact rates with susceptible hosts. Alternatively, negative individual-level effects on the amplification of different phage genotypes may be balanced by positive population-level effects on overall phage population growth. This could be tested using the matching-allele model of CRISPR-phage specificity presented in Chapter 3. Second, how do the impacts – if any - of relative levels of diversity depend on the pattern of host-pathogen specificity? Does host diversity beget pathogen diversity through increased specialisation, or are pathogens with broad

host ranges more likely to emerge when hosts are diverse? Simulations suggest that bacteria and phage can coevolve to form highly-diversified communities (Childs *et al.*, 2012), and the results I presented in Chapter 4 suggest that the ARD led to bacteria and phage diversification. However, other factors such as stochastic transmission chains (Antia *et al.*, 2003), or the balance between the availability of susceptible hosts and selection pressure from host diversity (Chabas *et al.*, 2018) could instead favour host range expansion at an individual-level, as I demonstrated in Chapter 3.

Concluding remarks

This thesis has examined how host diversity can shape bacteria-phage dynamics through both ecological and evolutionary effects. However, there is a need for further understanding of how the effects of immune diversity affects bacteria-phage dynamics in more realistic – and therefore more complex – environments. Life history trade-offs, alternative specificity networks, metapopulation dynamics, and phage diversity and evolution are just a small number of the possible other factors that can alter the tempo, mode, and course of bacteria-phage coevolution. Capturing the vast complexity of microbial communities requires a concerted combination of theoretical, experimental, and observational studies. Work that approaches such complexity is key to tackling the specific challenges posed by bacteria, in particular the treatment of bacterial infections. More refined computational, bioinformatic, and experimental techniques continually offer new ways to build bridges between the lab bench and a patient's bedside, or the computer screen and the geothermal pool.

Bacteria-phage interactions are also an ideal system to test more general ideas about host-pathogen dynamics. The work presented in this thesis aids our understanding of how host ecology has immediate epidemiological and long term evolutionary consequences, and provides evidence for synergy between them. Emergent infectious diseases such as SARS-CoV-2 illustrate the urgency of fundamental, “blue-skies” work to guide and compliment public health preparations and responses.

References

Acevedo-Whitehouse, K., Gulland, F., Greig, D., & Amos, W. (2003). Inbreeding: disease susceptibility in California sea lions. *Nature, 422*(6927), 35.

Agrawal, A., & Lively, C. M. (2002). Infection genetics: gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research, 4*(1), 79-90.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Selected papers of Hirotugu Akaike* (pp. 199-213). New York: Springer.

Alexander, H., & Day, T. (2010). Risk factors for the evolutionary emergence of pathogens. *Journal of The Royal Society Interface, 7*(51), 1455-1474.

Alseth, E. O., Pursey, E., Luján, A. M., McLeod, I., Rollie, C., & Westra, E. R. (2019). Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature, 574*(7779), 549-552.

Altermatt, F., & Ebert, D. (2008). Genetic diversity of Daphnia magna populations enhances resistance to parasites. *Ecology Letters, 11*(9), 918-928. Retrieved from http://dx.doi.org/10.1111/j.1461-0248.2008.01203.x. doi:10.1111/j.1461-0248.2008.01203.x

Anderson, R. M., & May, R. M. (1985). Vaccination and herd immunity to infectious diseases. *Nature, 318*(6044), 323-329.

Anderson, R. M., & May, R. M. (1992). *Infectious diseases of humans: dynamics and control*: Oxford university press.

Andersson, A. F., & Banfield, J. F. (2008). Virus population dynamics and acquired virus resistance in natural microbial communities. *Science, 320*(5879), 1047-1050. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18497291. doi:10.1126/science.1157358

Antia, R., Regoes, R. R., Koella, J. C., & Bergstrom, C. T. (2003). The role of evolution in the emergence of infectious diseases. *Nature, 426*(6967), 658.

Ashby, B., & King, K. C. (2015). Diversity and the maintenance of sex by parasites. *Journal of Evolutionary Biology, 28*(3), 511-520.

Baer, B., & Schmid-Hempel, P. (1999). Experimental variation in polyandry affects parasite loads and fitness in a bumble-bee. *Nature, 397*(6715), 151-154.

Baer, B., & Schmid-Hempel, P. (2001). Unexpected consequences of polyandry for parasitism and fitness in the bumblebee, Bombus terrestris. *Evolution, 55*(8), 1639-1643.

Baer, B., & Schmid-Hempel, P. (2003). Bumblebee workers from different sire groups vary in susceptibility to parasite infection. *Ecology Letters, 6*(2), 106-110.

Bates, D. M., M.; Bolker, B.; Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Beckett, S. J., & Williams, H. T. (2013). Coevolutionary diversification creates nested-modular structure in phage–bacteria interaction networks. *Interface focus, 3*(6), 20130033.

Benmayor, R., Hodgson, D. J., Perron, G. G., & Buckling, A. (2009). Host mixing and disease emergence. *Current Biology, 19*(9), 764-767.

Bohannan, B. J., & Lenski, R. E. (2000). Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecology Letters, 3*(4), 362-377.

Booth, R. (2020). BAME groups hit harder by Covid-19 than white people, UK study suggests. *The Guardian, 7*.

Borg, J., Kiær, L. P., Lecarpentier, C., Goldringer, I., Gauffreteau, A., Saint-Jean, S., . . . Enjalbert, J. (2018). Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research, 221*, 298-313.

Böttiger, M. (1987). A study of the sero-immunity that has protected the Swedish population against poliomyelitis for 25 years. *Scandinavian journal of infectious diseases, 19*(6), 595-601.

Brockhurst, M. A., Buckling, A., Poullain, V., & Hochberg, M. E. (2007). The impact of migration from parasite-free patches on antagonistic host-parasite coevolution. *Evolution: International Journal of Organic Evolution, 61*(5), 1238-1243.

Brockhurst, M. A., Chapman, T., King, K. C., Mank, J. E., Paterson, S., & Hurst, G. D. (2014). Running with the Red Queen: the role of biotic conflicts in evolution. *Proceedings of the Royal Society B: Biological Sciences, 281*(1797), 20141382.

Broniewski, J. M., Meaden, S., Paterson, S., Buckling, A., & Westra, E. R. (2020). The effect of phage genetic diversity on bacterial resistance evolution. *The ISME journal*, 1-9.

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., . . . van der Oost, J. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science, 321*(5891), 960-964. doi:10.1126/science.1159689

Brown, M. (2019). VCVglmm: Manipulating model outputs from MCMCglmm and lme4 objects, usually involving variance-covariance matrices (Version 0.0.0.9). Retrieved from https://github.com/Euphrasiologist/VCVglmm

Brundage, J. F., & Shanks, G. D. (2007). What really happened during the 1918 influenza pandemic? The importance of bacterial secondary infections. *The Journal of infectious diseases, 196*(11), 1717-1718.

Buckling, A., & Brockhurst, M. A. (2012). Bacteria-virus coevolution. *Advances in Experimental Medicine and Biology, 751*, 347-370. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22821466. doi:10.1007/978-1-4614-3567-9_16

Buckling, A., & Rainey, P. B. (2002a). Antagonistic coevolution between a bacterium and a bacteriophage. *Proceedings of the Royal Society of London B: Biological Sciences, 269*(1494), 931-936.

Buckling, A., & Rainey, P. B. (2002b). The role of parasites in sympatric and allopatric host diversification. *Nature, 420*(6915), 496-499.

Bull, R. A., Luciani, F., McElroy, K., Gaudieri, S., Pham, S. T., Chopra, A., . . . White, P. A. (2011). Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. *PLoS Pathogens, 7*(9), e1002243.

Burch, C. L., & Chao, L. (2000). Evolvability of an RNA virus is determined by its mutational neighbourhood. *Nature, 406*(6796), 625-628.

Burnham, K. P., & Anderson, D. R. (2003). *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261-304.

Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R., & O'Toole, G. A. (2012). The CRISPR/Cas Adaptive Immune System of *Pseudomonas aeruginosa* Mediates Resistance to Naturally Occurring and Engineered Phages. *Journal of Bacteriology, 194*(21), 5728-5738. Retrieved from http://jb.asm.org/content/194/21/5728.abstract. doi:10.1128/jb.01184-12

Cady, K. C., & O'Toole, G. A. (2011). Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. *Journal of bacteriology, 193*(14), 3433-3445.

Carte, J., Christopher, R. T., Smith, J. T., Olson, S., Barrangou, R., Moineau, S., . . . Terns, M. P. (2014). The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in S treptococcus thermophilus. *Molecular microbiology, 93*(1), 98-112.

Chabas, H., Lion, S., Nicot, A., Meaden, S., van Houte, S., Moineau, S., . . . Gandon, S. (2018). Evolutionary emergence of infectious diseases in heterogeneous host populations. *PLoS Biology, 16*(9), e2006738. Retrieved from https://doi.org/10.1371/journal.pbio.2006738. doi:10.1371/journal.pbio.2006738

Chabas, H., Nicot, A., Meaden, S., Westra, E. R., Tremblay, D. M., Pradier, L., . . . Gandon, S. (2019). Variability in the durability of CRISPR-Cas immunity. *Philosophical Transactions of the Royal Society B, 374*(1772), 20180097.

Chabas, H., van Houte, S., Høyland-Kroghsbo, N. M., Buckling, A., & Westra, E. R. (2016). Immigration of susceptible hosts triggers the evolution of alternative parasite defence strategies. *Proc. R. Soc. B, 283*(1837), 20160721.

Chaudhry, W. N., Pleška, M., Shah, N. N., Weiss, H., McCall, I. C., Meyer, J. R., . . . Levin, B. R. (2018). Leaky resistance and the conditions for the existence of lytic bacteriophage. *PLoS Biology, 16*(8), e2005971. Retrieved from https://doi.org/10.1371/journal.pbio.2005971. doi:10.1371/journal.pbio.2005971

Chevallereau, A., Meaden, S., Fradet, O., Landsberger, M., Maestri, A., Biswas, A., . . . Westra, E. R. (2020). Exploitation of the cooperative behaviors of anti-CRISPR phages. *Cell Host & Microbe, 27*(2), 189-198. e186.

Childs, L. M., England, W. E., Young, M. J., Weitz, J. S., & Whitaker, R. J. (2014). CRISPR-induced distributed immunity in microbial populations. *PloS One, 9*(7), e101710.

Childs, L. M., Held, N. L., Young, M. J., Whitaker, R. J., & Weitz, J. S. (2012). Multiscale model of CRISPR-induced coevolutionary dynamics: diversification at the interface of Lamarck and Darwin. *Evolution: International Journal of Organic Evolution, 66*(7), 2015-2029.

Chohan, J. K. (2017). Reclaiming the Food System: Agroecological Pedagogy and the IALA María Cano2. *Alternautas–Vol. 4–Issue 2–December 2017*, 13.

Chuang, J. H., & Li, H. (2004). Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome. *PLOS Biology, 2*(2).

Civitello, D. J., Cohen, J., Fatima, H., Halstead, N. T., Liriano, J., McMahon, T. A., . . . Young, S. (2015). Biodiversity inhibits parasites: broad evidence for the dilution effect. *Proceedings of the National Academy of Sciences, 112*(28), 8667-8671.

Common, J., Morley, D., van Houte, S., & Westra, E. R. (2019). CRISPR-Cas immunity leads to a coevolutionary arms race between *Streptococcus thermophilus* and lytic phage. *Philosophical Transactions of the Royal Society B: Biological Sciences*. doi:10.1098/rstb.2018.0098

Common, J., & Westra, E. R. (2019). CRISPR evolution and bacteriophage persistence in the context of population bottlenecks. *RNA Biology*.

Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K., & Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Communications, 3*, 945. Retrieved from http://dx.doi.org/10.1038/ncomms1937.

doi:http://www.nature.com/ncomms/journal/v3/n7/suppinfo/ncomms1937_S1.html

Dawkins, R. (1979). Twelve misunderstandings of kin selection. *Zeitschrift für Tierpsychologie, 51*(2), 184-200.

Dennehy, J. J. (2012). What can phages tell us about host-pathogen coevolution? *International journal of evolutionary biology, 2012*.

Dennehy, J. J., Friedenberg, N. A., Holt, R. D., & Turner, P. E. (2006). Viral ecology and the maintenance of novel host use. *The American Naturalist, 167*(3), 429-439.

Dennehy, J. J., Friedenberg, N. A., Yang, Y. W., & Turner, P. E. (2007). Virus population extinction via ecological traps. *Ecology Letters, 10*(3), 230-240.

Desai, S. D., & Currie, R. W. (2015). Genetic diversity within honey bee colonies affects pathogen load and relative virus levels in honey bees, Apis mellifera L. *Behavioral Ecology and Sociobiology, 69*(9), 1527-1541.

Deveau, H., Barrangou, R., Garneau, J. E., Labonté, J., Fremaux, C., Boyaval, P., . . . Moineau, S. (2008). Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *Journal of bacteriology, 190*(4), 1390-1400.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature, 418*(6898), 700-707.

Diekmann, O., & Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation* (Vol. 5): John Wiley & Sons.

Dobson, A. (2004). Population dynamics of pathogens with multiple host species. *The American Naturalist, 164*(S5), S64-S78.

Drake, J. W. (2009). Avoiding dangerous missense: thermophiles display especially low mutation rates. *PLoS Genetics, 5*(6), e1000520.

Duffy, S., Turner, P. E., & Burch, C. L. (2006). Pleiotropic costs of niche expansion in the RNA bacteriophage Φ6. *Genetics, 172*(2), 751-757.

Dyer, C. (2002). *Making a living in the middle ages: the people of Britain 850-1520*: Yale University Press.

Early, P., Huang, H., Davis, M., Calame, K., & Hood, L. (1980). An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell, 19*(4), 981-992.

Eastwood, J. R., Ribot, R. F., Rollins, L. A., Buchanan, K. L., Walder, K., Bennett, A. T., & Berg, M. L. (2017). Host heterozygosity and genotype rarity affect viral dynamics in an avian subspecies complex. *Scientific Reports, 7*(1), 13310.

Ekroth, A. K. E., Rafaluk-Mohr, C., & King, K. C. (2019). Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proceedings of the Royal Society B: Biological Sciences, 286*(1911), 20191811. Retrieved from https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2019.1811 %X There is evidence that human activities are reducing the population genetic diversity of species worldwide. Given the prediction that parasites better exploit genetically homogeneous host populations, many species could be vulnerable to disease outbreaks. While agricultural studies have shown the devastating effects of infectious disease in crop monocultures, the widespread

nature of this diversity–disease relationship remains unclear in natural systems. Here, we provide broad support that high population genetic diversity can protect against infectious disease by conducting a meta-analysis of 23 studies, with a total of 67 effect sizes. We found that parasite functional group (micro- or macroparasite) affects the presence of the effect and study setting (field or laboratory-based environment) influences the magnitude. Our study also suggests that host genetic diversity is overall a robust defence against infection regardless of host reproduction, parasite host range, parasite diversity, virulence and the method by which parasite success was recorded. Combined, these results highlight the importance of monitoring declines of host population genetic diversity as shifts in parasite distributions could have devastating effects on at-risk populations in nature. doi:doi:10.1098/rspb.2019.1811

Elton, C. S. (1958). *The ecology of invasions by animals and plants*. London: Methuen & Co.

Ferrieres, L., Hémery, G., Nham, T., Guérout, A.-M., Mazel, D., Beloin, C., & Ghigo, J.-M. (2010). Silent mischief: bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *Journal of Bacteriology, 192*(24), 6418-6427.

Fine, P. E. (1993). Herd immunity: history, theory, practice. *Epidemiologic reviews, 15*(2), 265-302.

Fine, P. E., Eames, K., & Heymann, D. L. (2011). “Herd immunity”: a rough guide. *Clinical infectious diseases, 52*(7), 911-916.

Flanagan, J., Brodie, E., Weng, L., Lynch, S., Garcia, O., Brown, R., . . . Wiener-Kronish, J. (2007). Loss of bacterial diversity during antibiotic treatment of intubated patients colonized with *Pseudomonas aeruginosa*. *Journal of Clinical Microbiology, 45*(6), 1954-1962.

Flores, C. O., Valverde, S., & Weitz, J. S. (2013). Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *The ISME journal, 7*(3), 520-532.

Fortuna, M. A., Barbour, M. A., Zaman, L., Hall, A. R., Buckling, A., & Bascompte, J. (2019). Coevolutionary dynamics shape the structure of bacteria-phage infection networks. *Evolution, 73*(5), 1001-1011.

Fox, J. P., Elveback, L., Scott, W., GATEWOOD, L., & Ackerman, E. (1971). Herd immunity: basic concept and relevance to public health immunization practices. *American Journal of Epidemiology, 94*(3), 179-189.

Frankel, O. H. (1939). Analytical yield investigations on New Zealand wheat: IV. Blending varieties of wheat. *The Journal of Agricultural Science, 29*(2), 249-261.

Gaba, S., & Ebert, D. (2009). Time-shift experiments as a tool to study antagonistic coevolution. *Trends in Ecology & Evolution, 24*(4), 226-232.

Gandon, S. (2004). Evolution of multihost parasites. *Evolution, 58*(3), 455-469.

Gandon, S., Buckling, A., Decaestecker, E., & Day, T. (2008). Host–parasite coevolution and patterns of adaptation across time and space. *Journal of evolutionary biology, 21*(6), 1861-1866.

Gandon, S., Hochberg, M. E., Holt, R. D., & Day, T. (2013). What limits the evolutionary emergence of pathogens? *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1610). doi:10.1098/rstb.2012.0086

Gandon, S., & Michalakis, Y. (2002). Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *Journal of evolutionary biology, 15*(3), 451-462.

Ganz, H. H., & Ebert, D. (2010). Benefits of host genetic diversity for resistance to infection depend on parasite diversity. *Ecology, 91*(5), 1263-1268. Retrieved from http://dx.doi.org/10.1890/09-1243.1. doi:10.1890/09-1243.1

Garneau, J. E., Dupuis, M.-E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., . . . Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature, 468*(7320), 67-71. Retrieved from http://dx.doi.org/10.1038/nature09523. doi:10.1038/nature09523

Gijsbers, E. F., Schuitemaker, H., & Kootstra, N. A. (2012). HIV-1 transmission and viral adaptation to the host. *Future Virology, 7*(1), 63-71.

Gokhale, C. S., Papkou, A., Traulsen, A., & Schulenburg, H. (2013). Lotka–Volterra dynamics kills the Red Queen: population size fluctuations and associated stochasticity dramatically change host-parasite coevolution. *BMC Evolutionary Biology, 13*(1), 254.

Gomulkiewicz, R., Thompson, J. N., Holt, R. D., Nuismer, S. L., & Hochberg, M. E. (2000). Hot spots, cold spots, and the geographic mosaic theory of coevolution. *The American Naturalist, 156*(2), 156-174.

Gorter, F. A., Hall, A. R., Buckling, A., & Scanlan, P. D. (2015). Parasite host range and the evolution of host resistance. *Journal of evolutionary biology, 28*(5), 1119-1130.

Gould, G. M. (1966). *Anomalies and curiosities of medicine*: Blacksleet River.

Grissa, I., Vergnaud, G., & Pourcel, C. (2007a). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC bioinformatics, 8*(1), 172.

Grissa, I., Vergnaud, G., & Pourcel, C. (2007b). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research, 35*(suppl_2), W52-W57.

Gutiérrez, S., Michalakis, Y., & Blanc, S. (2012). Virus population bottlenecks during within-host progression and host-to-host transmission. *Current Opinion in Virology, 2*(5), 546-555.

Guyader, S., & Burch, C. L. (2008). Optimal foraging predicts the ecology but not the evolution of host specialization in bacteriophages. *PLoS ONE, 3*(4), e1946.

Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software, 33*(2), 1-22. Retrieved from http://www.jstatsoft.org/v33/i02/.

Haldane, J. B. S. (1949). Suggestions as to quantitative measurement of rates of evolution. *Evolution, 3*(1), 51-56.

Hall, A. R., Scanlan, P. D., & Buckling, A. (2011a). Bacteria-phage coevolution and the emergence of generalist pathogens. *The American Naturalist, 177*(1), 44-53.

Hall, A. R., Scanlan, P. D., Morgan, A. D., & Buckling, A. (2011b). Host–parasite coevolutionary arms races give way to fluctuating selection. *Ecology Letters, 14*(7), 635-642.

Hamer, W. H. (1906). *Epidemic disease in England: the evidence of variability and of persistency of type*: Bedford Press.

Harbeck, M., Seifert, L., Hänsch, S., Wagner, D. M., Birdsell, D., Parise, K. L., . . . Keim, P. (2013). Yersinia pestis DNA from skeletal remains from the 6th century AD reveals insights into Justinianic Plague. *PLoS pathogens, 9*(5).

Hedrich, A. (1933). MONTHLY ESTIMATES OF THE CHILD POPULATION "SUSCEPTIBLE'TO MEASLES, 1900–1931, BALTIMORE, MD. *American Journal of Epidemiology, 17*(3), 613-636.

Heesterbeek, J. A. P. (2002). A brief history of R 0 and a recipe for its calculation. *Acta biotheoretica, 50*(3), 189-204.

Heineman, R. H., Springman, R., & Bull, J. J. (2008). Optimal foraging by bacteriophages through host avoidance. *The American Naturalist, 171*(4), E149-E157.

Hesse, E., Best, A., Boots, M., Hall, A. R., & Buckling, A. (2015). Spatial heterogeneity lowers rather than increases host–parasite specialization. *Journal of evolutionary biology, 28*(9), 1682-1690.

Hesse, E., & Buckling, A. (2016). Host population bottlenecks drive parasite extinction during antagonistic coevolution. *Evolution, 70*(1), 235-240.

Heymann, D., & Aylward, B. (2008). Mass vaccination in public health. *Control of communicable diseases manual. 19th ed. Washington, DC: American Public Health Association*.

Honour, H., & Fleming, J. (2005). *A world history of art*: Laurence King Publishing.

Horvath, P., Coute-Monvoisin, A. C., Romero, D. A., Boyaval, P., Fremaux, C., & Barrangou, R. (2009). Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *International Journal of Food Microbiology, 131*(1), 62-70. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18635282. doi:10.1016/j.ijfoodmicro.2008.05.030

Horvath, P., Romero, D. A., Coute-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., . . . Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *Journal of Bacteriology, 190*(4), 1401-1412. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18065539. doi:10.1128/JB.01415-07

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346-363.

Hudson, P. J., Dobson, A. P., & Lafferty, K. D. (2006). Is a healthy ecosystem one that is rich in parasites? *Trends in Ecology & Evolution, 21*(7), 381-385.

Hudson, P. J., Dobson, A. P., & Newborn, D. (1998). Prevention of population cycles by parasite removal. *Science, 282*(5397), 2256-2258.

Hyman, P., & Abedon, S. T. (2010). Bacteriophage host range and bacterial resistance. *Advances in applied microbiology, 70*, 217-248.

Iranzo, J., Lobkovsky, A. E., Wolf, Y. I., & Koonin, E. V. (2013). Evolutionary Dynamics of the Prokaryotic Adaptive Immunity System CRISPR-Cas in an Explicit Ecological Context. *Journal of Bacteriology, 195*(17), 3834-3844. Retrieved from http://jb.asm.org/content/195/17/3834.abstract. doi:10.1128/jb.00412-13

Jackson, S. A., McKenzie, R. E., Fagerlund, R. D., Kieper, S. N., Fineran, P. C., & Brouns, S. J. (2017). CRISPR-Cas: Adapting to change. *Science, 356*(6333), eaal5056.

Jiang, N., Weinstein, J. A., Penland, L., White, R. A., Fisher, D. S., & Quake, S. R. (2011). Determinism and stochasticity during maturation of the zebrafish antibody repertoire. *Proceedings of the National Academy of Sciences, 108*(13), 5348-5353.

Jiang, W., Maniv, I., Arain, F., Wang, Y., Levin, B. R., & Marraffini, L. A. (2013). Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genetics, 9*(9), e1003844. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/24086164. doi:10.1371/journal.pgen.1003844

John, T. J., & Samuel, R. (2000). Herd immunity and herd effect: new insights and definitions. *European journal of epidemiology, 16*(7), 601-606.

Jończyk, E., Kłak, M., Międzybrodzki, R., & Górski, A. (2011). The influence of external factors on bacteriophages. *Folia Microbiologica, 56*(3), 191-200.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Duran, C. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics, 28*(12), 1647-1649.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., . . . Mitchell, C. E. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature, 468*(7324), 647-652.

Keesing, F., Holt, R. D., & Ostfeld, R. S. (2006). Effects of species diversity on disease risk. *Ecology Letters, 9*(4), 485-498.

Kiær, L. P., Skovgaard, I. M., & Østergård, H. (2009). Grain yield increase in cereal variety mixtures: a meta-analysis of field trials. *Field Crops Research, 114*(3), 361-373.

King, K. C., Delph, L. F., Jokela, J., & Lively, C. M. (2011). Coevolutionary hotspots and coldspots for host sex and parasite local adaptation in a snail–trematode interaction. *Oikos, 120*(9), 1335-1340. Retrieved from http://dx.doi.org/10.1111/j.1600-0706.2011.19241.x. doi:10.1111/j.1600-0706.2011.19241.x

Koonin, E. V., Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology, 37*, 67-78. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28605718. doi:10.1016/j.mib.2017.05.008

Koskella, B. (2014). Bacteria-phage interactions across time and space: merging local adaptation and time-shift experiments to understand phage evolution.

Am Nat, 184 Suppl 1, S9-21. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25061680. doi:10.1086/676888

Koskella, B., & Brockhurst, M. A. (2014). Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiology Reviews, 38*(5), 916-931.

Koskella, B., & Lively, C. M. (2007). Advice of the rose: experimental coevolution of a trematode parasite and its snail host. *Evolution, 61*(1), 152-159.

Koskella, B., & Meaden, S. (2013). Understanding bacteriophage specificity in natural microbial communities. *Viruses, 5*(3), 806-823.

Laanto, E., Hoikkala, V., Ravantti, J., & Sundberg, L.-R. (2017). Long-term genomic coevolution of host-parasite interaction in the natural environment. *Nature Communications, 8*(1), 111.

Landsberger, M., Gandon, S., Meaden, S., Rollie, C., Chevallereau, A., Chabas, H., . . . van Houte, S. (2018). Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. *Cell, 174*(4), 908-916. e912.

Laughton, R. (2017). *A Matter Of Scale: A study of the productivity, financial viability and multifunctional benefits of small farms (20 ha and less)*. Retrieved from Landworkers Alliance:

Lawton, J. H. (1999). Are there general laws in ecology? *Oikos*, 177-192.

Lenski, R. E. (1991). Quantifying fitness and gene stability in microorganisms. *Biotechnology, 15*, 173-192.

Lenski, R. E., & Levin, B. R. (1985). Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. *The American Naturalist, 125*(4), 585-602.

Lévesque, C., Duplessis, M., Labonté, J., Labrie, S., Fremaux, C., Tremblay, D., & Moineau, S. (2005). Genomic organization and molecular analysis of virulent bacteriophage 2972 infecting an exopolysaccharide-producing Streptococcus thermophilus strain. *Applied and Environmental Microbiology, 71*(7), 4057-4068.

Levin, B. R., Moineau, S., Bushman, M., & Barrangou, R. (2013). The population and evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. *PLoS Genet, 9*(3), e1003312. doi:10.1371/journal.pgen.1003312

Littman, R. J. (2009). The plague of Athens: epidemiology and paleopathology. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine: A Journal of Translational and Personalized Medicine, 76*(5), 456-467.

Lively, C. M. (2010). The Effect of Host Genetic Diversity on Disease Spread. *The American Naturalist, 175*(6), E149-E152. Retrieved from http://dx.doi.org/10.1086/652430. doi:10.1086/652430

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., & Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature, 438*(7066), 355-359.

López-Uribe, M. M., Appler, R. H., Youngsteadt, E., Dunn, R. R., Frank, S. D., & Tarpy, D. R. (2017). Higher immunocompetence is associated with higher

genetic diversity in feral honey bee colonies (Apis mellifera). *Conservation Genetics, 18*(3), 659-666.

Luijckx, P., Fienberg, H., Duneau, D., & Ebert, D. (2013). A Matching-Allele Model Explains Host Resistance to Parasites. *Current Biology*.

MacArthur, R. H., & Pianka, E. R. (1966). On optimal use of a patchy environment. *The American Naturalist, 100*(916), 603-609.

Machida-Hirano, R. (2015). Diversity of potato genetic resources. *Breeding science, 65*(1), 26-40.

Magurran, A. E., & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature, 422*(6933), 714-716.

Mann, C. C. (2005). *1491: New revelations of the Americas before Columbus*: Alfred a Knopf Incorporated.

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science, 322*(5909), 1843-1845.

Marston, M. F., Pierciey, F. J., Shepard, A., Gearin, G., Qi, J., Yandava, C., . . . Martiny, J. B. (2012). Rapid diversification of coevolving marine Synechococcus and a virus. *Proceedings of the National Academy of Sciences, 109*(12), 4544-4549.

Martínez-García, E., Aparicio, T., de Lorenzo, V., & Nikel, P. I. (2014). New transposon tools tailored for metabolic engineering of Gram-negative microbial cell factories. *Frontiers in Bioengineering and Biotechnology, 2*, 46.

Martiny, J. B., Riemann, L., Marston, M. F., & Middelboe, M. (2014). Antagonistic coevolution of marine planktonic viruses and their hosts. *Annual review of marine science, 6*, 393-414.

May, R. M., Gupta, S., & McLean, A. R. (2001). Infectious disease dynamics: what characterizes a successful invader? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 356*(1410), 901-910.

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., . . . He, F. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters, 10*(10), 995-1015.

Meagher, S. (1999). Genetic diversity and Capillaria hepatica (Nematoda) prevalence in Michigan deer mouse populations. *Evolution, 53*(4), 1318-1324.

Meyer, J. R., Dobias, D. T., Weitz, J. S., Barrick, J. E., Quick, R. T., & Lenski, R. E. (2012). Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science, 335*(6067), 428-432.

Middelboe, M., Holmfeldt, K., Riemann, L., Nybroe, O., & Haaber, J. (2009). Bacteriophages drive strain diversification in a marine Flavobacterium: implications for phage resistance and physiological properties. *Environmental microbiology, 11*(8), 1971-1982.

Mitchell, C. E., Tilman, D., & Groth, J. V. (2002). Effects of grassland plant species diversity, abundance, and composition on foliar fungal disease. *Ecology, 83*(6), 1713-1726.

Mizoguchi, K., Morita, M., Fischer, C. R., Yoichi, M., Tanji, Y., & Unno, H. (2003). Coevolution of bacteriophage PP01 and Escherichia coli O157: H7 in continuous culture. *Applied and environmental microbiology, 69*(1), 170-176.

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology, 155*(Pt 3), 733-740. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19246744. doi:10.1099/mic.0.023960-0

Moore, R. I. (2008). *The formation of a persecuting society: authority and deviance in Western Europe 950-1250*: John Wiley & Sons.

Morens, D. M., & Fauci, A. S. (2007). The 1918 influenza pandemic: insights for the 21st century. *The Journal of infectious diseases, 195*(7), 1018-1028.

Morgan, A. D., Gandon, S., & Buckling, A. (2005). The effect of migration on local adaptation in a coevolving host–parasite system. *Nature, 437*(7056), 253-256.

Morley, D., Broniewski, J. M., Westra, E. R., Buckling, A., & van Houte, S. (2017). Host diversity limits the evolution of parasite local adaptation. *Molecular Ecology, 26*(7), 1756-1763.

Morran, L. T., Schmidt, O. G., Gelarden, I. A., Parrish, R. C., & Lively, C. M. (2011). Running with the Red Queen: host-parasite coevolution selects for biparental sex. *Science, 333*(6039), 216-218.

Mundt, C. C. (2002). Use of multiline cultivars and cultivar mixtures for disease management. *Annual review of phytopathology, 40*(1), 381-410.

Nirenberg, D. (2015). *Communities of violence: Persecution of minorities in the middle ages-updated edition*: Princeton University Press.

O'Brien, S. J., Roelke, M. E., Marker, L., Newman, A., Winkler, C., Meltzer, D., . . . Wildt, D. E. (1985). Genetic basis for species vulnerability in the cheetah. *Science, 227*(4693), 1428-1434.

Ohtsuki, A., & Sasaki, A. (2006). Epidemiology and disease-control under gene-for-gene plant–pathogen interaction. *Journal of Theoretical Biology, 238*(4), 780-794.

Ostfeld, R. S., & Keesing, F. (2012). Effects of host diversity on infectious disease. *Annual Review of Ecology, Evolution, and Systematics, 43*.

Pachepsky, E., Crawford, J. W., Bown, J. L., & Squire, G. (2001). Towards a general theory of biodiversity. *Nature, 410*(6831), 923.

Paez-Espino, D., Morovic, W., Sun, C. L., Thomas, B. C., Ueda, K.-i., Stahl, B., . . . Banfield, J. F. (2013). Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nature Communications, 4*, 1430.

Paez-Espino, D., Sharon, I., Morovic, W., Stahl, B., Thomas, B. C., Barrangou, R., & Banfield, J. F. (2015). CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *MBio, 6*(2). Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25900652. doi:10.1128/mBio.00262-15

Paigen, K., & Petkov, P. (2010). Mammalian recombination hot spots: properties, control and evolution. *Nature Reviews Genetics, 11*(3), 221-233.

Papagrigorakis, M. J., Yapijakis, C., & Synodinos, P. N. (2008). Typhoid fever epidemic in ancient Athens. In *Paleomicrobiology* (pp. 161-173): Springer.

Payne, P., Geyrhofer, L., Barton, N. H., & Bollback, J. P. (2018). CRISPR-based herd immunity can limit phage epidemics in bacterial populations. *eLife, 7*, e32035.

Pearman, P. B., & Garner, T. W. (2005). Susceptibility of Italian agile frog populations to an emerging strain of Ranavirus parallels population genetic diversity. *Ecology Letters, 8*(4), 401-408.

Poisot, T., Lounnas, M., & Hochberg, M. E. (2013). The structure of natural microbial enemy-victim networks. *Ecological Processes, 2*(1), 1-9.

Poullain, V., Gandon, S., Brockhurst, M. A., Buckling, A., & Hochberg, M. E. (2008). The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. *Evolution, 62*(1), 1-11.

Pyenson, N. C., & Marraffini, L. A. (2020). Co-evolution within structured bacterial communities results in multiple expansion of CRISPR loci and enhanced immunity. *eLife, 9*.

Pyke, G. H., Pulliam, H. R., & Charnov, E. L. (1977). Optimal foraging: a selective review of theory and tests. *The quarterly review of biology, 52*(2), 137-154.

Quigley, B. J., López, D. G., Buckling, A., McKane, A. J., & Brown, S. P. (2012). The mode of host–parasite interaction shapes coevolutionary dynamics and the fate of host cooperation. *Proc. R. Soc. B, 279*(1743), 3742–3748.

R Core Team. (2018). R: A language and environment for statistical computing (Version 3.5.2 "Eggshell Igloo"). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rast, J. P., & Litman, G. W. (1994). T-cell receptor gene homologs are present in the most primitive jawed vertebrates. *Proceedings of the National Academy of Sciences, 91*(20), 9248-9252.

Reber, A., Castella, G., Christe, P., & Chapuisat, M. (2008). Experimentally increased group diversity improves disease resistance in an ant species. *Ecology Letters, 11*(7), 682-689.

Rego, R. O., Bestor, A., Štefka, J., & Rosa, P. A. (2014). Population bottlenecks during the infectious cycle of the Lyme disease spirochete *Borrelia burgdorferi*. *PloS One, 9*(6), e101009.

Reichert, T. A., Sugaya, N., Fedson, D. S., Glezen, W. P., Simonsen, L., & Tashiro, M. (2001). The Japanese experience with vaccinating schoolchildren against influenza. *New England Journal of Medicine, 344*(12), 889-896.

Reiss, E. R., & Drinkwater, L. E. (2018). Cultivar mixtures: a meta-analysis of the effect of intraspecific diversity on crop yield. *Ecological Applications, 28*(1), 62-77.

Reluga, T., Meza, R., Walton, D. B., & Galvani, A. P. (2007). Reservoir interactions and disease emergence. *Theoretical Population Biology, 72*(3), 400-408.

Richter, C., Dy, R. L., McKenzie, R. E., Watson, B. N., Taylor, C., Chang, J. T., . . . Fineran, P. C. (2014). Priming in the Type IF CRISPR-Cas system triggers

strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Research, 42*(13), 8516-8526.

Roscher, C., Schumacher, J., Foitzik, O., & Schulze, E.-D. (2007). Resistance to rust fungi in *Lolium perenne* depends on within-species variation and performance of the host species in grasslands of different plant diversity. *Oecologia, 153*(1), 173-183.

Rosen, W. (2010). *Justinian's flea: plague, empire and the birth of Europe*: Random House.

Sasaki, A. (2000). Host-parasite coevolution in a multilocus gene-for-gene system. *Proceedings of the Royal Society B: Biological Sciences, 267*(1458), 2183-2188.

Scanlan, P. D., Hall, A. R., Burlinson, P., Preston, G., & Buckling, A. (2013). No effect of host–parasite co-evolution on host range expansion. *Journal of evolutionary biology, 26*(1), 205-209.

Schmidt, A. M., Linksvayer, T. A., Boomsma, J. J., & Pedersen, J. S. (2011). No benefit in diversity? The effect of genetic variation on survival and disease resistance in a polygynous social insect. *Ecological Entomology, 36*(6), 751-759.

Schrag, S., & Mittler, J. (1996). Host-parasite coexistence: the role of spatial refuges in stabilizing bacteria-phage interactions. *The American Naturalist, 148*(2), 348-377.

Semenova, E., Jore, M. M., Datsenko, K. A., Semenova, A., Westra, E. R., Wanner, B., . . . Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proceedings of the National Academy of Sciences, 108*(25), 10098-10103. Retrieved from http://www.pnas.org/content/108/25/10098.abstract. doi:10.1073/pnas.1104144108

Sharp, P. M., & Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harbor perspectives in medicine, 1*(1), a006841.

Smithson, J. B., & Lenne, J. M. (1996). Varietal mixtures: a viable strategy for sustainable productivity in subsistence agriculture. *Annals of Applied Biology, 128*(1), 127-158.

Stephens, D. S. (2008). Vaccines for the unvaccinated: protecting the herd. In: The University of Chicago Press.

Sun, C. L., Thomas, B. C., Barrangou, R., & Banfield, J. F. (2016). Metagenomic reconstructions of bacterial CRISPR loci constrain population histories. *ISME J, 10*(4), 858-870. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26394009. doi:10.1038/ismej.2015.162

Swarts, D. C., Mosterd, C., Van Passel, M. W., & Brouns, S. J. (2012). CRISPR interference directs strand specific spacer acquisition. *PloS One, 7*(4), e35888.

Telfer, S., Lambin, X., Birtles, R., Beldomenico, P., Burthe, S., Paterson, S., & Begon, M. (2010). Species interactions in a parasite community drive infection risk in a wildlife population. *Science, 330*(6001), 243-246.

Therneau, T. M., & Lumley, T. (2015). Package ‘survival’. *R Top Doc, 128*.

Thompson, J. N. (2005). *The geographic mosaic of coevolution*: University of Chicago Press.

Thorne, E. T., & Williams, E. S. (1988). Disease and endangered species: the black-footed ferret as a recent example. *Conservation Biology, 2*(1), 66-74.

Thucydides, Warner, R., & Finley, M. (1972). *History of the Peloponnesian War*.

Topley, W., & Wilson, G. (1923). The spread of bacterial infection. The problem of herd-immunity. *Epidemiology & Infection, 21*(3), 243-249.

Torsvik, V., & Øvreås, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology, 5*(3), 240-245.

Travisano, M., & Lenski, R. E. (1996). Long-term experimental evolution in Escherichia coli. IV. Targets of selection and the specificity of adaptation. *Genetics, 143*(1), 15-26.

Tripathi, A., Tripathi, D. K., Chauhan, D., Kumar, N., & Singh, G. (2016). Paradigms of climate change impacts on some major food sources of the world: a review on current knowledge and future prospects. *Agriculture, ecosystems & environment, 216*, 356-373.

Turchin, P. (2018). *Historical dynamics: Why states rise and fall* (Vol. 26): Princeton University Press.

Vale, P. F., Lafforgue, G., Gatchitch, F., Gardan, R., Moineau, S., & Gandon, S. (2015). Costs of CRISPR-Cas-mediated resistance in Streptococcus thermophilus. *Proc. R. Soc. B, 282*(1812), 20151270.

Vale, P. F., & Little, T. J. (2010). CRISPR-mediated phage resistance and the ghost of coevolution past. *Proceedings of the Royal Society of London B: Biological Sciences, 277*(1691), 2097-2103.

van Houte, S., Buckling, A., & Westra, E. R. (2016a). Evolutionary Ecology of Prokaryotic Immune Mechanisms. *Microbiology and Molecular Biology Reviews, 80*(3), 745-763. Retrieved from http://mmbr.asm.org/content/80/3/745.abstract. doi:10.1128/mmbr.00011-16

van Houte, S., Ekroth, A. K. E., Broniewski, J. M., Chabas, H., Ashby, B., Bondy-Denomy, J., . . . Westra, E. R. (2016b). The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature, 532*(7599), 385-388. Retrieved from http://dx.doi.org/10.1038/nature17436. doi:10.1038/nature17436

Verberk, W. C., Van Der Velde, G., & Esselink, H. (2010). Explaining abundance–occupancy relationships in specialists and generalists: a case study on aquatic macroinvertebrates in standing waters. *Journal of Animal Ecology, 79*(3), 589-601.

Weinberger, A. D., Wolf, Y. I., Lobkovsky, A. E., Gilmore, M. S., & Koonin, E. V. (2012). Viral diversity threshold for adaptive immunity in prokaryotes. *MBio, 3*(6), e00456-00412.

Weinstein, J. A., Jiang, N., White, R. A., Fisher, D. S., & Quake, S. R. (2009). High-throughput sequencing of the zebrafish antibody repertoire. *Science, 324*(5928), 807-810.

Weissman, J. L., Holmes, R., Barrangou, R., Moineau, S., Fagan, W. F., Levin, B., & Johnson, P. L. (2018a). Immune loss as a driver of coexistence during host-phage coevolution. *The ISME Journal, 12*(2), 585.

Weissman, J. L., Laljani, R., Fagan, W. F., & Johnson, P. L. (2018b). Ecology Shapes Microbial Immune Strategy: Temperature and Oxygen as Determinants of the Incidence of CRISPR Adaptive Immunity. *bioRxiv*, 326330.

Weitz, J. S., Poisot, T., Meyer, J. R., Flores, C. O., Valverde, S., Sullivan, M. B., & Hochberg, M. E. (2013). Phage–bacteria infection networks. *Trends in microbiology, 21*(2), 82-91.

Westra, E. R., Semenova, E., Datsenko, K. A., Jackson, R. N., Wiedenheft, B., Severinov, K., & Brouns, S. J. (2013). Type IE CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genetics, 9*(9), e1003742.

Westra, E. R., Sunderhauf, D., Landsberger, M., & Buckling, A. (2017). Mechanisms and consequences of diversity-generating immune strategies. *Nature Reviews: Immunology, 17*(11), 719-728. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28787398. doi:10.1038/nri.2017.78

Westra, E. R., Swarts, D. C., Staals, R. H., Jore, M. M., Brouns, S. J., & van der Oost, J. (2012a). The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. *Annual Review of Genetics, 46*, 311-339.

Westra, E. R., van Erp, P. B., Künne, T., Wong, S. P., Staals, R. H., Seegers, C. L., . . . Severinov, K. (2012b). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Molecular cell, 46*(5), 595-605.

Westra, E. R., van Houte, S., Oyesiku-Blakemore, S., Makin, B., Broniewski, Jenny M., Best, A., . . . Buckling, A. (2015). Parasite Exposure Drives Selective Evolution of Constitutive versus Inducible Defense. *Current Biology, 25*(8), 1043-1049. Retrieved from http://dx.doi.org/10.1016/j.cub.2015.01.065. doi:10.1016/j.cub.2015.01.065

Whiteman, N. K., Kimball, R. T., & Parker, P. G. (2007). Co-phylogeography and comparative population genetics of the threatened Galápagos hawk and three ectoparasite species: ecology shapes population histories within parasite communities. *Molecular Ecology, 16*(22), 4759-4773.

Wichman, H. A., Millstein, J., & Bull, J. J. (2005). Adaptive molecular evolution for 13,000 phage generations: a possible arms race. *Genetics, 170*(1), 19-31.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York. Retrieved from http://ggplot2.org

Woolhouse, M. E., Dye, C., Etard, J.-F., Smith, T., Charlwood, J., Garnett, G., . . . Quinnell, R. (1997). Heterogeneities in the transmission of infectious agents:

implications for the design of control programs. *Proceedings of the National Academy of Sciences, 94*(1), 338-342.

World Health Organisation. (2020). HIV/AIDS. Retrieved from https://www.who.int/gho/hiv/en/

Wu, D., Wu, T., Liu, Q., & Yang, Z. (2020). The SARS-CoV-2 outbreak: what we know. *International Journal of Infectious Diseases*.

Yates, A., Antia, R., & Regoes, R. R. (2006). How do pathogen evolution and host heterogeneity interact in disease emergence? *Proceedings of the Royal Society B: Biological Sciences, 273*(1605), 3075-3083.

Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H., & O'Toole, G. A. (2009). Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *Journal of Bacteriology, 191*(1), 210-219.

Zhu, Y., Chen, H., Fan, J., & Wang, Y. (2000). Genetic diversity and disease control in rice. *Nature, 406*(6797), 718.

Appendix to Chapter 2 – Supplementary Material

Comparison		Hazard Ratio	Standard Error	*z*	*p*
ΔCRISPR	CRISPR				
10^{-2}	10^{-2}	0.35	2.31	-1.26	1.00
10^{-3}	10^{-3}	0.24	2.12	-1.93	0.84
10^{-4}	10^{-4}	0.19	2.31	-1.98	0.81
10^{-5}	10^{-4}	1.46	2.03	0.53	1.00
10^{-6}	10^{-6}	0.10	2.31	-2.80	0.24
10^{-7}	10^{-7}	1.31	2.03	0.39	1.00
10^{-8}	10^{-8}	4.36	2.11	1.97	0.82
10^{-9}	10^{-9}	2.68	2.08	1.34	0.99

Table A2.1 Tukey's significant difference tests of phage extinction hazard ratios between bottleneck treatments in ΔCRISPR and CRISPR background

	CRISPR		Sensitive		SM	
Bottleneck	Mean	95% CIs	Mean	95% CIs	Mean	95% CIs
10^{-2}	0.92	0.87–0.96	0.04	0.02–0.08	0.03	0.01–0.07
10^{-3}	0.88	0.82–0.93	0.03	0.01–0.07	0.08	0.05–0.14
10^{-4}	0.76	0.69–0.83	0.05	0.02–0.09	0.19	0.13–0.26
10^{-5}	0.82	0.75–0.88	0.08	0.05–0.14	0.10	0.06–0.15
10^{-6}	0.38	0.27–0.49	0.61	0.50–0.72	0.01	0.00–0.06
10^{-7}	0.24	0.17–0.31	0.73	0.65–0.80	0.03	0.01–0.07
10^{-8}	0.13	0.08–0.19	0.88	0.81–0.92	0.00	0.00–0.00
10^{-9}	0.13	0.08–0.20	0.87	0.80–0.92	0.00	0.00–0.00

Table A2.2 Mean relative frequency and 95% CIs of CRISPR, sensitive and SM immune phenotypes in the bacterial culture at 3 d.p.i. when both the host and phage are bottlenecked

Bottleneck	CRISPR Mean	CRISPR 95% CIs	Sensitive Mean	Sensitive 95% CIs	SM Mean	SM 95% CIs
10^{-2}	0.85	0.79–0.91	0.03	0.01–0.07	0.11	0.07–0.17
10^{-3}	0.81	0.74–0.86	0.05	0.02–0.09	0.15	0.09–0.20
10^{-4}	0.75	0.68–0.82	0.08	0.04–0.13	0.17	0.12–0.24
10^{-5}	0.78	0.71–0.84	0.08	0.05–0.14	0.14	0.09–0.20
10^{-6}	0.74	0.66–0.80	0.13	0.08–0.18	0.13	0.09–0.18
10^{-7}	0.60	0.52–0.68	0.24	0.18–0.32	0.16	0.11–0.20
10^{-8}	1.00	1.00–1.00	0.00	0.00–0.00	0.00	0.00–0.00
10^{-9}	NA	NA	NA	NA	NA	NA

Table A2.3 Mean relative frequency and 95% CIs of CRISPR, sensitive and SM immune phenotypes in the bacterial culture at 3 d.p.i. when host was bottlenecked and phage was supplemented at each transfer

Bottleneck	CRISPR Mean	CRISPR 95% CIs	Sensitive Mean	Sensitive 95% CIs	SM Mean	SM 95% CIs
10^{-4} bottleneck	0.76	0.69–0.82	0.05	0.00–0.09	0.19	0.13–0.26
Small dilution	0.56	0.44–0.67	0.39	0.28–0.50	0.06	0.02–0.12
10^{-6} bottleneck	0.37	0.27–0.49	0.61	0.49–0.72	0.01	0.00–0.06
Large dilution	0.08	0.01–0.0.32	0.92	0.68–0.99	0.00	0.00–0.00

Table A2.4 Mean relative frequency and 95% CIs of CRISPR, sensitive and SM immune phenotypes in the bacterial culture at 3 d.p.i. when host was either bottlenecked or comparably diluted

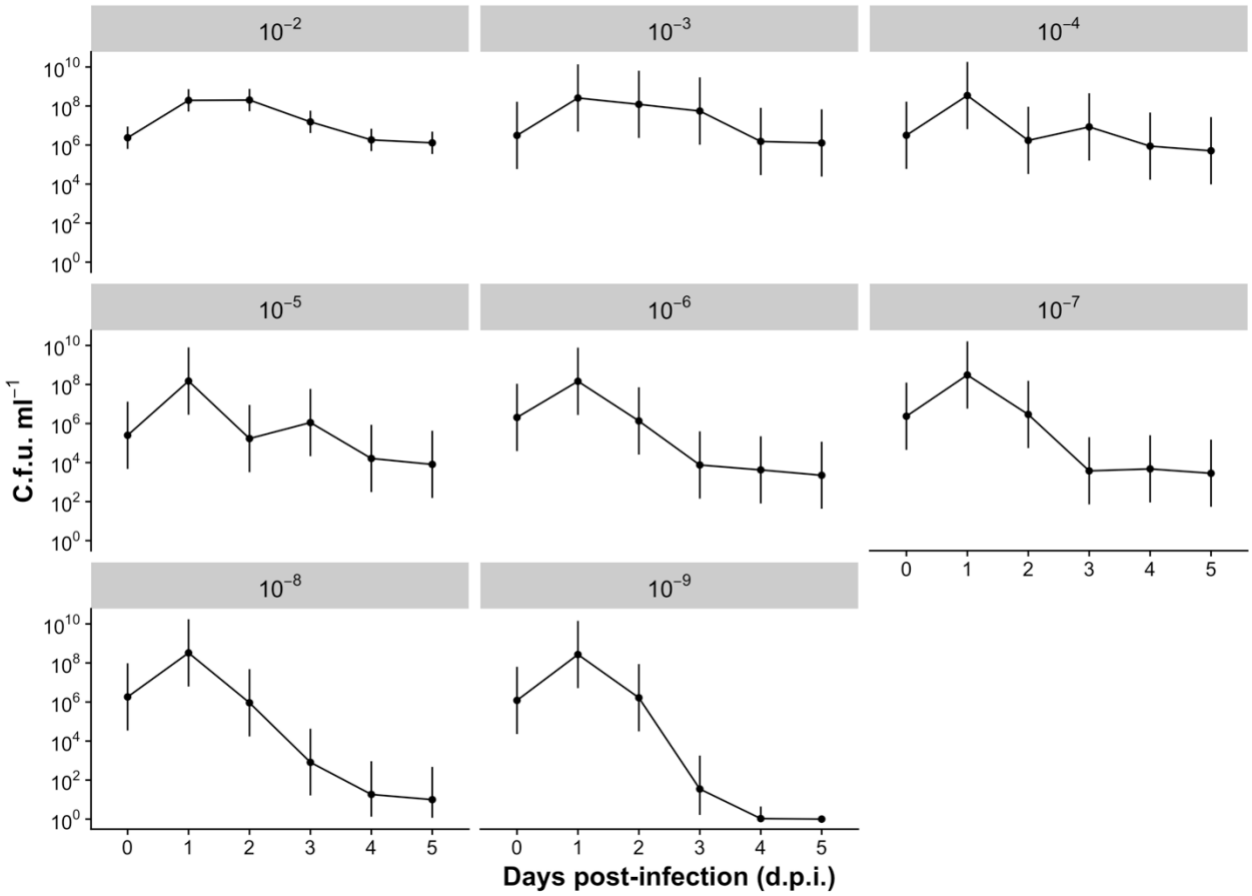


Figure A2.1 Host population dynamics when WT *P. aeruginosa* PA14 host is bottlenecked in the absence of phage (phage-negative control). Mean colony-forming units (c.f.u.) ml⁻are shown for different bottleneck treatments (ranging from 10^{-2}-10^{-9} dilutions at each transfer, as indicated above each panel). The detection limit is 200 c.f.u. ml^{-1}. Error bars correspond to 95% confidence intervals (CIs). N = 6 for all treatments.

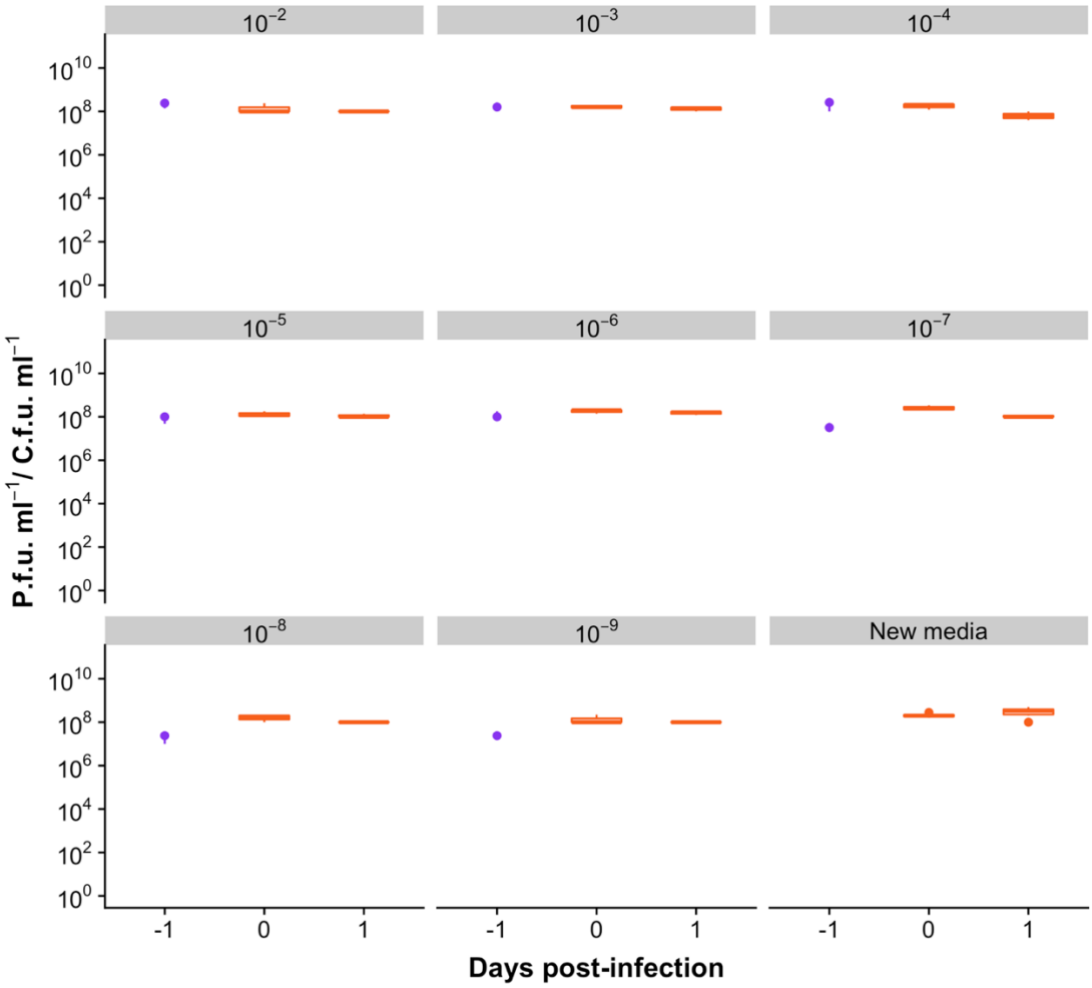


Figure A1.2 Boxplots of DMS3vir titre (plaque-forming units (pfu) ml^{-1}) after overnight incubation in used and new media (*N*=6). Used media was generated by inoculating fresh 6ml of M9 minimal media (supplemented with 0.2% glucose) from 10^{-2}-10^{-9} from an overnight culture of WT *P. aeruginosa* PA14, followed by overnight incubation at 37˚ and 180rpm. Cultures were then sampled to measure host density (purple points and bars are the median, minimum and maximum colony-forming units (cfu) ml^{-1}). To remove bacterial cells and generate preparations of used media, 5ml of each culture was then centrifuged at 3500rpm for 25 mins, and filtered through 0.45µm filters into clean glass vials. 10^8 pfu ml^{-1} DMS3vir was then added to each preparation, and also to fresh (new) media. To measure phage titre, samples of the

(Figure A1.2 caption contd.)

preparations then underwent a chloroform extraction followed by spot assays on a top lawn of *P. aeruginosa* PA14 *csy3::lacZ*. The preparations were then incubated overnight at 37˚ and 180rpm. The next day, phage titres were again measured using chloroform extraction followed by spot assays.

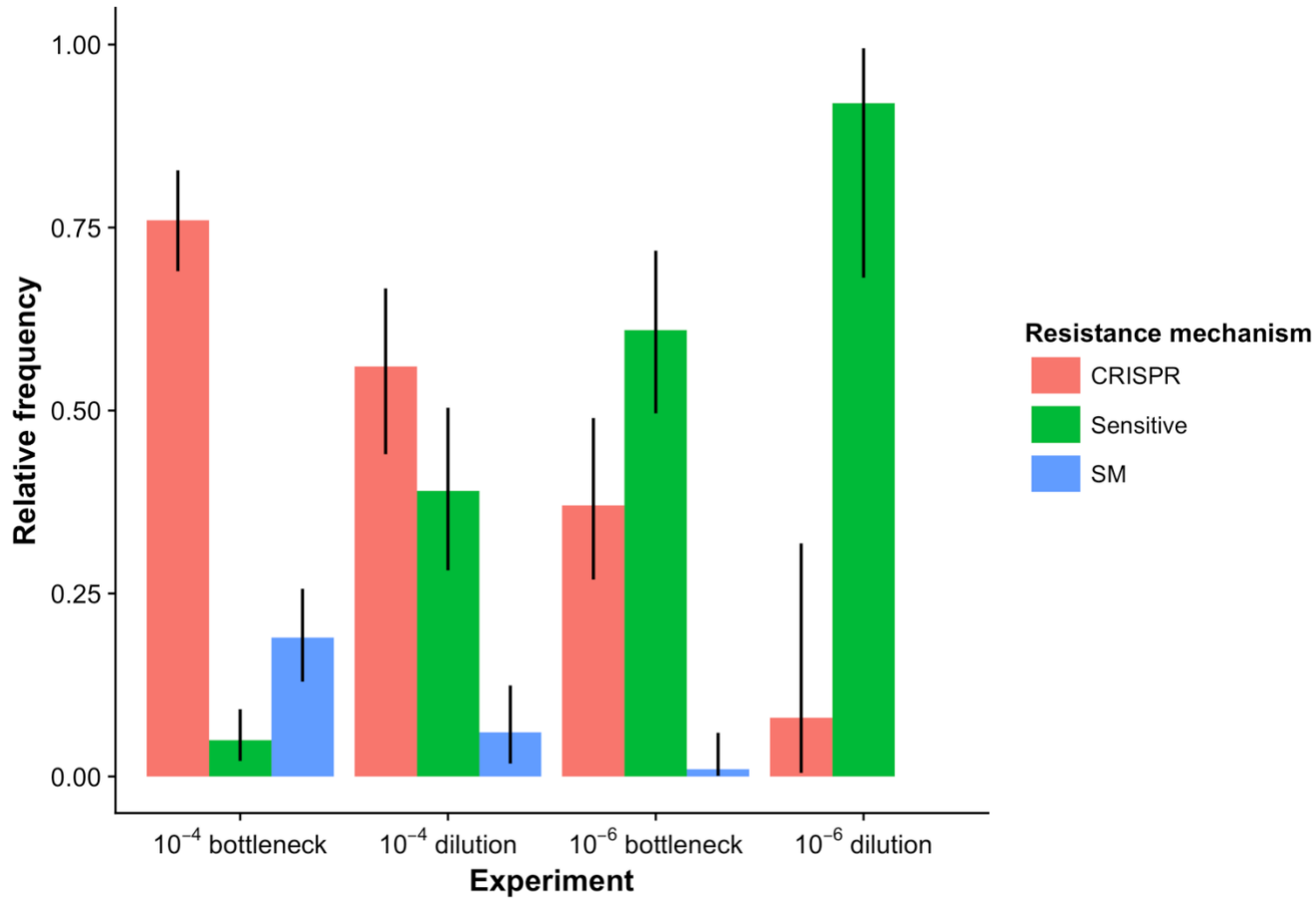


Figure A2.3 Mean relative frequencies of bacterial clones with CRISPR immunity, surface mutation (SM) resistance, or sensitive phenotypes, at 3 days post-infection. A 10^{-4} or 10^{-6} bottleneck corresponds to either 0.6µl or 0.006µl culture, respectively, in 6ml of fresh media at each transfer. A 10^{-4} dilution corresponds to 0.6µl culture in 6 ml of fresh media at each transfer, and a 10^{-6} dilution corresponds to 0.6µl culture in 600ml of fresh media at each transfer (see Materials & Methods). Error bars correspond to 95% confidence intervals (CIs). N=6 for all treatments

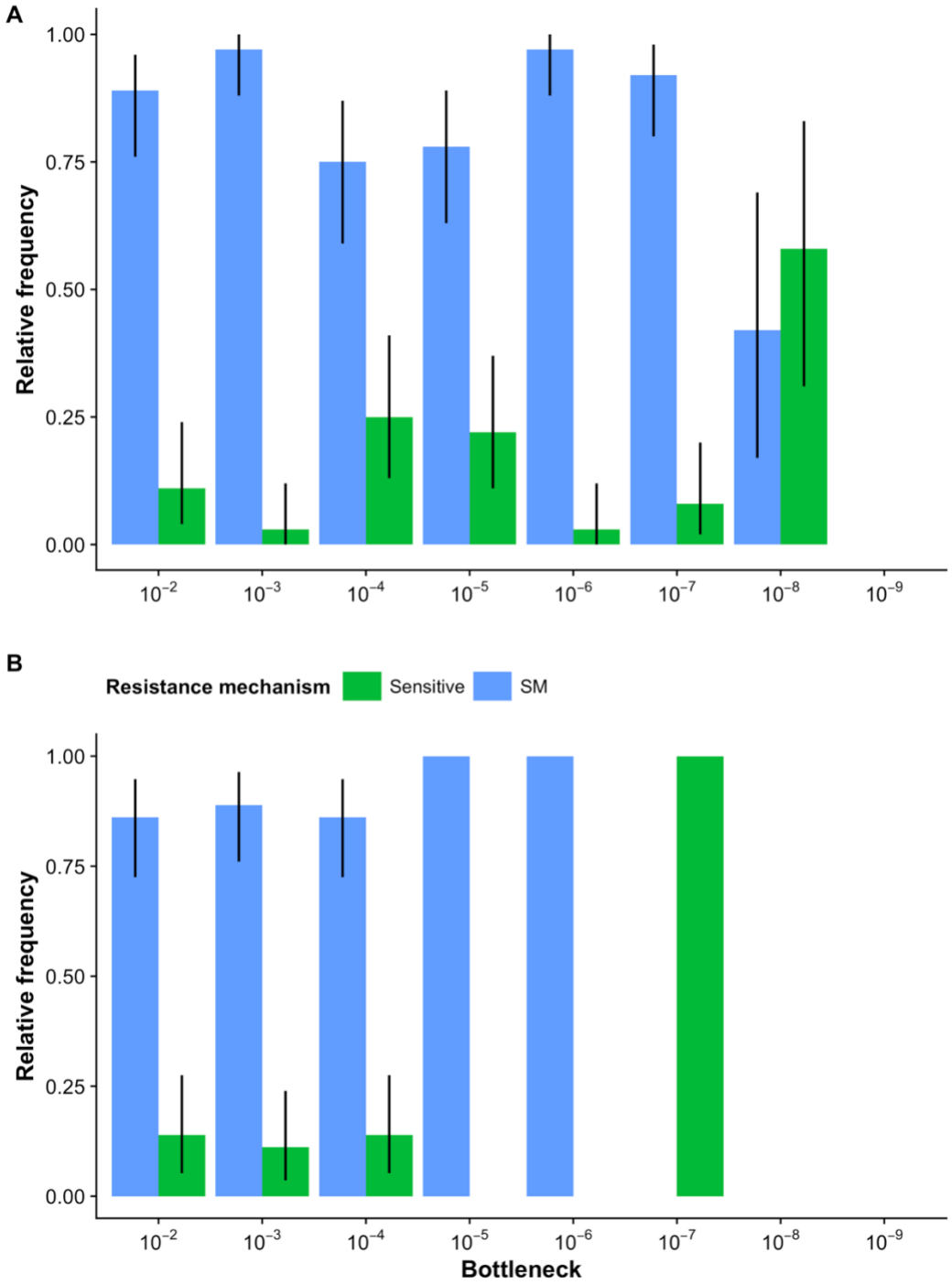


Figure A2.4 Mean relative frequencies of bacterial clones with CRISPR immunity, surface mutation (SM) resistance, or sensitive phenotypes, at 3 days post-infection in a ΔCRISPR background **A)** When both host and phage were bottlenecked and **B)** when host was bottlenecked and phage supplemented at each transfer. Error bars correspond to 95% confidence intervals (CIs) N=6 for all treatments.

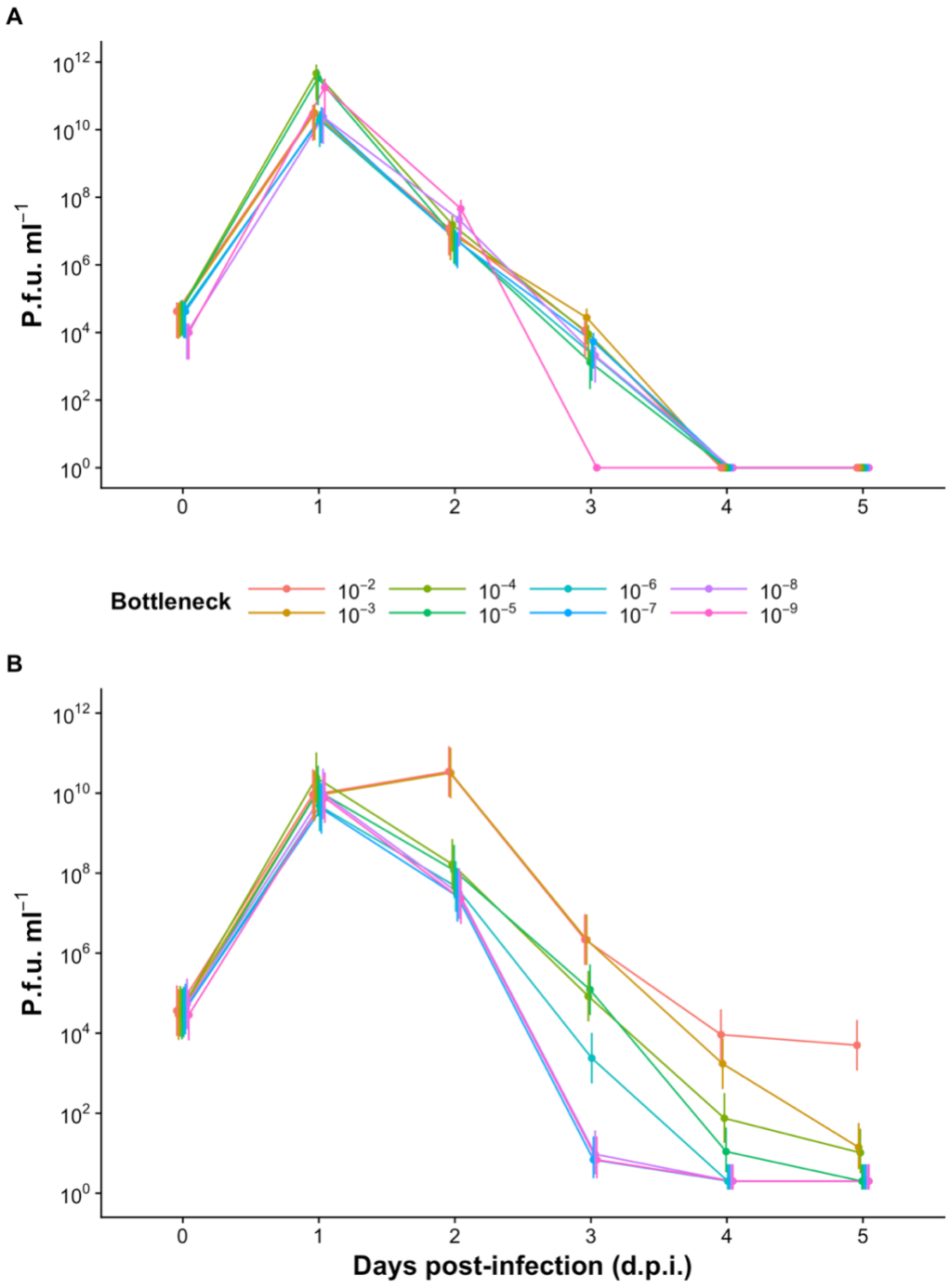


Figure A2.5 Phage dynamics when both only the host is bottlenecked in **A)** CRISPR background and **B)** ΔCRISPR control. Mean plaque-forming units (p.f.u.) ml^{-1} are shown. Values of 10^0 represent phage titres of zero, rather than one. Error bars correspond to 95% confidence intervals (CIs). N=6 for all treatments.

Appendix to Chapter 3 - Supplementary Materials & Methods

Library of BIMs and escape phages

11 *P. aeruginosa* PA14 bacteriophage-insensitive mutants (BIMs) that were known to have a single CRISPR2 spacer were selected from the collection of clones used in van Houte *et al.* (2016b). The additional 13 BIMs were generated by evolving *P. aeruginosa* PA14 in the presence of DMS3vir. 6ml of M9 minimal media (supplemented with 0.2% glucose; M9m) was inoculated with approximately 10^6 colony-forming units (cfu) of WT *P. aeruginosa* and 10^4 plaque-forming units (pfu) of phage in glass vials. After 24hrs, samples from the infection were plated on LB agar. Potential CRISPR clones were identified through phenotypic and PCR analyses as described previously (Westra 2015; van Houte 2016). CRISPR amplicon sequencing (SourceBioscience, UK) confirmed that each spacer carried by a BIM was unique, so that all clones used in downstream experiments carried a different spacer. Spacer sequences were mapped against the DMS3vir genome (Genbank accession: NC_008717.1) using Geneious v9.1.8 (Kearse *et al.*, 2012) to confirm that spacers did not target overlapping regions of the phage genome. See Table S1 in Supporting Information for the spacer sequences of each BIM.

To generate 24 phage clones that could infect each BIM (escape phage), 15ml LB was inoculated with approximately 10^6 cfu of a single BIM and approximately 10^6 pfu DMS3vir. We also added approximately 10^6 of *P. aeruginosa* PA14 *csy3::lacZ* to provide a pool of sensitive hosts on which phage could replicate and hence supply novel escape mutations. These cultures were then incubated overnight at 37˚C and

180rpm. Escape phage from these amplifications were identified by spot assay on a top lawn of the BIM with which they were originally mixed, and were then plaque-purified to ensure a monoclonal phage stock. Each escape phage was challenged against the entire BIM library to check for a one-to-one infection match. A successful infection was defined if a clear lysis zone was visible in the top lawn of the target BIM.

Generating labelled BIMs

The BIMs chosen for transformation were such that a single clone could be monitored in each of the 3-clone mixtures (that is, BIMs 1, 4, 7, 10, 13, 16, 19, and 22; see Table S1), which enabled us to measure relative frequency and fitness of a labelled BIM through time by performing a blue:white screen when plating on LB agar supplemented with 40µg/ml X-gal.

All cloning reactions to generate the labelled BIMs were carried out according to manufacturers' instructions unless stated otherwise. Restriction enzymes, Antarctic phosphatase, and T4 DNA ligase were purchased from NEB; High-Fidelity (HF) versions were used if available. Strains, primers, and plasmids used for molecular work are outlined in Table S2. We used the synthetic mini-Tn5 transposon vector pBAMD1-6 (Martínez-García *et al.*, 2014) to deliver the *lacZ* gene to target BIMs. pBAMD1-6 is a non-replicative vector in *P. aeruginosa* encoding a Tn5 transposase, which allows for insertion of a gentamicin resistance gene (GmR) as well as any cargo genes into the bacterial chromosome. To introduce *lacZ* as a cargo gene, we amplified it from PA14 *csy3::lacZ* using primers lacZ_amp_fw and lacZ_amp_rv

(Table S2) using Phusion High-Fidelity Polymerase (ThermoFisher). The PCR product was cleaned up (QIAgen PCR cleanup kit) and sub-cloned into pMA-RQ_Cas (Walker-Sünderhauf, unpublished) using NcoI-HF and KpnI-HF to generate a construct in which *lacZ* gene expression is driven by a constitutive β-lactamase promoter P3 (Genbank accession: J01749, region 4156..4233). Using standard molecular cloning protocols and restriction enzymes HindIII-HF and KpnI-HF, this promoter and the downstream *lacZ* gene was inserted into pBAMD1-6 to generate pBAM1(Gm)_lacZ. pBAM1(Gm)_lacZ was transferred into *E. coli* MFD*pir* by electroporation.

Tn5 insertions of the recipient BIMs were carried out by conjugative pBAM1(Gm)_lacZ delivery. *E. coli MFD*pir + pBAM1(Gm)_lacZ was used as donor and grown overnight in 5ml LB + 0.3mM diaminopimelic acid (DAP) + 30 µg/ml gentamicin at 37˚C, 180 rpm. Recipient BIMs were grown overnight in 5ml LB at 37˚C, 180rpm. 10ml of fresh media was inoculated from these overnight cultures, and grown at 37˚C and 250rpm until OD_{600} ~ 0.6, then pelleted and washed twice in 1x M9 salts, and resuspended in 1ml 1 x M9 salts. 600µl of donors were mixed with 200µl recipients, pelleted, and resuspended to a volume of 100µl. The entire donor-recipient mixture was pipetted onto sterile 0.2µm microfiber glass filters (Whatman) on LB agar + 0.3mM DAP plates and incubated for 2 days at 37˚C. To recover cells, filters were placed into 2.5ml LB and vortexed. 100µl of recovered cells were plated onto LB agar + 30 µg/ml gentamicin + 40µg/ml X-gal + 0.1mM IPTG plates and incubated at 37˚C for 2 days to select for BIMs with Tn5 insertions in their genome (absence of DAP selects against the donor strain).

Because Tn5 inserts at random positions in the *P. aeruginosa* genome, this may affect fitness. We therefore sampled three blue colonies of each transformed BIM and conducted 24hr competition experiments against their untransformed counterpart to verify their fitness was unaffected. The relative fitness of the transformed BIM was calculated as described previously (W_n = [(fraction transformant at t_n) * (1 – (fraction stain transformant at t_0))] / [(fraction transformant at t_0) * (1 – (fraction transformant at t_n)])(Westra *et al.*, 2015). If Tn5 insertion disrupted the CRISPR-Cas system, the transformed BIM would regain susceptibility to ancestral DMS3vir. We therefore checked for this by spotting ancestral DMS3vir on a top lawn of the transformed BIM. If no clear lysis zone was visible on the top lawn, we determined that the CRISPR-Cas system was functional. These checks confirmed that transformation with pBAM1(Gm)_lacZ did not affect relative fitness of the BIMs compared to their untransformed counterparts, or disrupt the CRISPR-Cas system.

Supplementary Figures

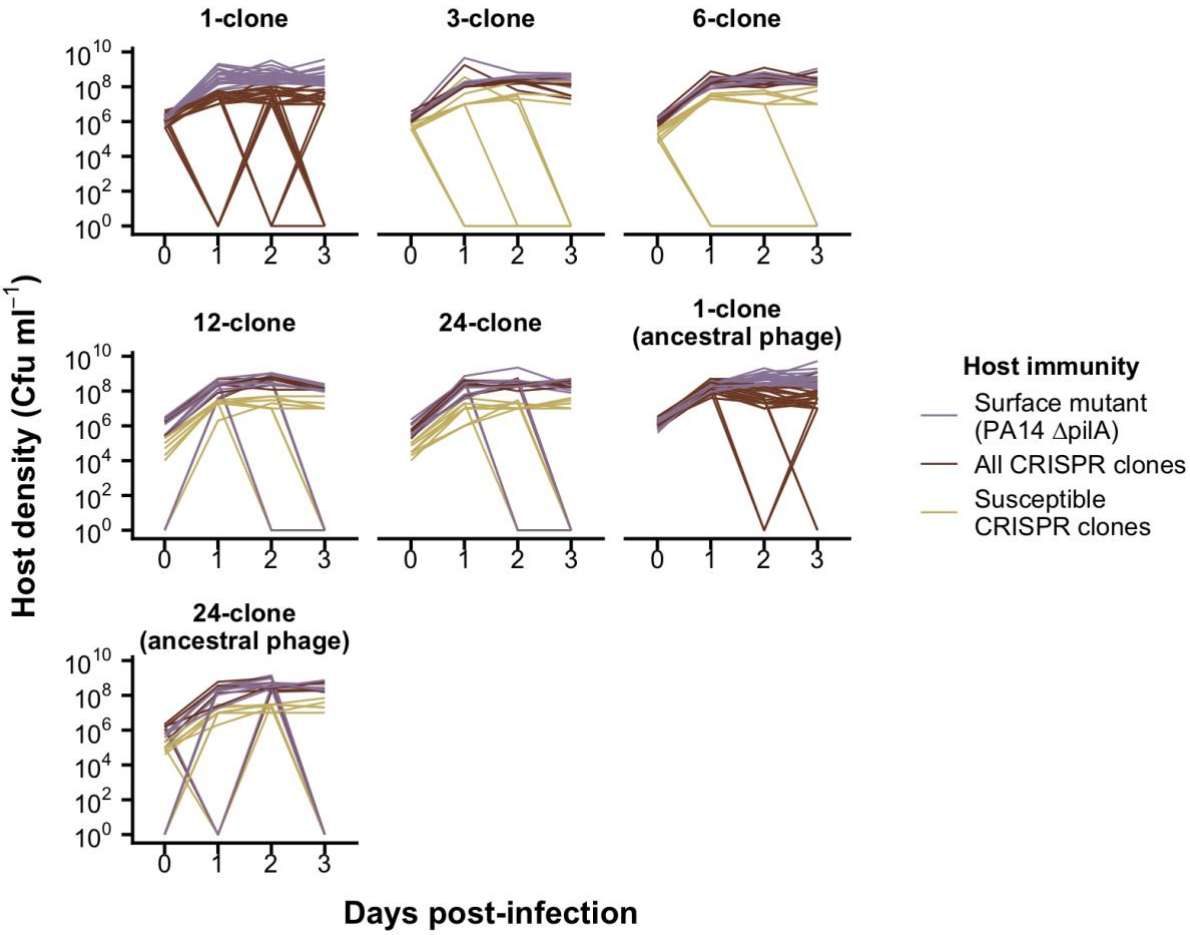


Figure A3.1 Density of surface mutants, all CRISPR clones, and the susceptible CRISPR clone

Population dynamics of the bacterial host population at different levels of CRISPR allele diversity in the host population (panels). Lines show the density expressed as colony-forming units (cfu) ml^{-1} in individual replicates at each day post-infection (x-axis). Line colours indicate host immunity: *P. aeruginosa* PA14 Δ*pilA* surface mutant, which is completely resistant to phage; all CRISPR clones from the BIM library; and susceptible CRISPR clones.

Supplementary Tables

Table A3.1 Sequence of spacers in the CRISPR2 locus of each of the 24 bacteriophage-insensitive mutants (BIMs) used in the co-culture experiment. Clones transformed to carry a *lacZ* reporter gene using pBAM1(Gm)_lacZ are highlighted in blue.

BIM	Spacer sequence
1	ATTTCAGTCCTTCCTGATCGCGTAGAGCCAAG
2	CATCTTCCCGCTCGATGGCGGTCAGCGTGCGC
3	CGCGTGAATGGCCCGGCGCTGAGCTGCGCTAT
4	AAGGGCATCAACCTGGCCGAAGGCGGCGCGCC
5	CGGTCGAACACGCCCTTATAGCGCTTCAGGCC
6	GATGTTCATCGCTGCCGGGCAGCGCGACATAC
7	AAACAGCGTCATGTCCAGGAGCTGCCGCTCGC
8	ACGGCAAGTTGAGTCTGGCCCTGGATGCTGAC
9	CCGGAAGTCCCGGCCGGTGTAGACGAGATAAA
10	GGCTCGACCAGGCGGCCCAGGGCGGCGTCGAT
11	TCAGGACCCCGACCAGATGGCGGCCGAAATGT
12	CGCCTGGAGACCCTGAAGGCCAATACCGAAAA
13	CCGAACGCATANANGGCGCANGGCACAGGGGT
14	ATGGGGATTCAGAGCTACGGCGATACCGCCCT
15	GAAATCGGCACCGCCACGAACCACCAGAACCT
16	GTCCAGCAGGATGCCGGCATCATCAACGAAAT
17	GGCAACGATCCCCACGAGCGGCTTTGGCACCT
18	CTCAACTCCGGCGCCGAAGACGTGATTGTCGA
19	GCGGGATCGCGGAGATAGCAGCTACGCTCGTA
20	ACTTTCACGACGACCCAGAAGCGTCGGCCGTT

Table A3.1 contd.

BIM	Spacer sequence
21	GCGGCAGGAGCGGCAGCGGGCGGCGGCAGTT
22	GCGATCAGNTGCGGCCAATCCGTGGACTGGGT
23	GATGGCGTCAAACTCGGCCTCCAGGCGCAGCG
24	AACCTCGCGCAGTCGTTGTCCAGCGGCATCAT

Table A3.2 Plasmids, Primers, and Strains used for molecular cloning work.

Plasmids			
Plasmid	**Reference**	**Genbank accession number**	**Culture conditions**
pMA-RQ_Cas	Walker-Sünderhauf (unpublished)		100 µg/mL Ampicillin, 50 µg/mL Gentamicin, 0.3 mM diaminopimelic acid. Needs a *pir* strain to replicate.
pBAMD1-6	Martinez-Garcia *et al.* 2014	KM403115	
pBAM1(Gm)_lacZ	This study		
Primers			
Primer	**Sequence (5' → 3')**		**Usage**
lacZ_amp_fw	TTACCATGGATGATTACGGAT TCACTGGCCGTCGT		Amplification of *lacZ* from PA14 *csy3:lacZ*. Adds NcoI and KpnI restriction sites onto amplicon.
lacZ_amp_rv	CAGGTACCTTATTTTTGACAC CAGACCAACTGGTAATGGT		

Table A3.2 contd.

Bacterial Strains		
Strain	**Reference/Supplier**	**Usage**
Pseudonomas aeruginosa PA14 *csy3::lacZ*	Zegans *et al.* 2009	Template for *lacZ* amplification.
E. coli CC18λpir	NEB	Cloning of promoter + *lacZ* onto pBAM1(Gm)
E. coli MFD*pir*	Ferrieres *et al.* 2010	Donor strain for pBAM1(Gm)_lacZ delivery
P. aeruginosa PA14 BIMs	This study; Table A3.1	Recipients for pBAM1(Gm)_lacZ delivery

Table A3.3 Primers used to amplify and sequence the protospacers of interest of phage that were shown to have undergone host shift (lost infectivity to the original clone and could only infect a new clone) from the phenotypic assay. The first column indicates if the protospacer was the original pre-evolved or the new protospacer, with the identity of the BIM the phage could infect shown in brackets. The primer sequence and binding direction are shown, and if the primer was used for PCR or sequencing reactions.

Phage	Sequence (5'- 3')	Bind direction	Primer
Original (7)	CCTGGACCTTCGCGCCGGAC	F	PCR
	GAGGTGAGGTCTTCGCTTTC	R	
	GTCGCACGGAATGTTCAGCGAG	R	Sequencing
Original (13)	TCTGGCCAGGCGCTCACAAACAA	F	PCR
	GAGCGGCTTTGGCACCTGGAAC	R	
	CCAAGTGTCGCTGCCGATCA	R	Sequencing
New (10)	AGCTGTCCACTGCGCTGGAC	F	PCR
	CCGGAACAGATGATCCCGTT	R	
	AATGTCAGCGCGGCGGTTGC	R	Sequencing
New (21)	CAGCGGCATCATGGGGCTGTTTG	F	PCR
	AGGTACTGAAGTTTTTGGAGGG	R	
	CCGCTGCTATCCAGACGGCC	F	Sequencing

Table A3.4 Protospacer sequences of evolved phage clones which showed host shift according to the phenotypic assay from replicate 3 of the 24-clone treatment at 1 day post-infection (dpi). The CRISPR-targeted protospacer and PAM sequences of the ancestral (WT) phage and of the pre-evolved phage are shown. Numbers 1-12 are independent phage isolates from the treatment, replicate and timepoint of interest. The second column indicates if the protospacer was the original pre-evolved or the new protospacer, with the identity of the protospacer shown in brackets. Protospacer-adjacent motif (PAM) and protospacer sequences are shown separately. SNPs and deletions, relative to WT DMS3vir, are highlighted in red.

24-clone, replicate 3, 1 dpi			
Phage	Protospacer	PAM sequence	Protospacer sequence
WT DMS3vir	Original (7)	GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
Pre-evolved protospacer 7		**A**G	CGCTCGCCGTCGAGGACCTGTACTGCGACAA**C**
1		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
2		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
3		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
4		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
5		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
6		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
7		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
8		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
9		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
10		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
11		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA
12		GG	CGCTCGCCGTCGAGGACCTGTACTGCGACAAA

Table A3.4 contd.

24-clone, replicate 3, 1 dpi			
Phage	Protospacer	PAM sequence	Protospacer sequence
WT DMS3vir	New (10)	GG	TAGCTGCGGCGGGACCCGGCGGACCAGCTCGG
Pre-evolved protospacer 7		GG	TAGCTGCGGCGGGACCCGGCGGACCAGCTCGG
Pre-evolved protospacer 10		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
1		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
2		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
3		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
4		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
5		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
6		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
7		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
8		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
9		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
10		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
11		GG	**C**AGCTGCGGCGGGACCCGGCGGACCAGCTCGG
12		GG	**C**ANCTGCGGCGGGACCCGGCGGACCAGCTCGG

Table A3.5 Protospacer sequences of evolved phage clones which showed host shift according to the phenotypic assay from replicate 5 of the 24-clone treatment at 2 days post-infection (dpi). The CRISPR-targeted protospacer and PAM sequences of the ancestral (WT) phage and of the pre-evolved phage are shown. Numbers 1-8 are independent phage isolates from the treatment, replicate and timepoint of interest. The second column indicates if the protospacer was the original pre-evolved or the new protospacer, with the identity of the protospacer shown in brackets. Protospacer-adjacent motif (PAM) and protospacer sequences are shown separately. SNPs and deletions, relative to WT DMS3vir, are highlighted in red.

24-clone, replicate 5, 2 dpi			
Phage	Protospacer	PAM sequence	Protospacer sequence
WT DMS3vir	Original (13)	GG	TGGGGACACGGGACGCGGTAGATACGCAAGCC
Pre-evolved protospacer 13		G**A**	TGGGGACACGGGACGCGGTAGATACGCAAGCC
1		GG	TGGGGACACGGGACGCGGTAGATACGCAAGCC
2		GG	TGGGGACACGGGACGCGGTAGATACGCAAGCC
3		GG	TGGGGACACGGGACGCGGTAGATACGCAAGCC
4		GG	TGGGGACACGGGACGCGGTAGATACGCAAGCC
5		GG	TGGGGACACGGGACGCGGTAGATACGCAAGCC
6		GG	TGGGGACACGGGACNCGGTAGATACGC**T**AGCC
7		GG	TGGGGACACGGGACGCGGTAGATACGCAAGCN
8		GG	TGGGGACACGGGACGNGGTAGATACGCAAGCC

Table A3.5 contd.

24-clone, replicate 5, 2 dpi			
Phage	Protospacer	PAM sequence	Protospacer sequence
WT DMS3vir	New (21)	GG	TGCGGCAGGAGCGGCAGCGGGCGGCGGCAGTT
Pre-evolved protospacer 13		GG	TGCGGCAGGAGCGGCAGCGGGCGGCGGCAGTT
Pre-evolved protospacer 21		GG	TGCGGCAG----------CGGGCGGNGGCAGTT
1		GG	T----------GCGGCAGCGGGCGGCGGCAGTT
2		GG	T----------GCGGCAGCGGGCGGCGGCAGTT
3		GG	T----------GCGGCAGCGGGCGGCGGCAGTT
4		GG	T----------GCGGCAGCGGGCGGCGGCAGTT
5		GG	T----------GCGGCAGCGGGCGGCGGCAGTT
6		GG	T----------GCGGCAGCGGGCGGCGGCAGTT
7		GG	T----------GCGGCAGCGGGCGGCGGCAGTT
8		GG	T----------GCGGCAGCGGGCGGCGGCAGTT

Appendix to Chapter 4 – Supplementary Materials

Supplementary Figures

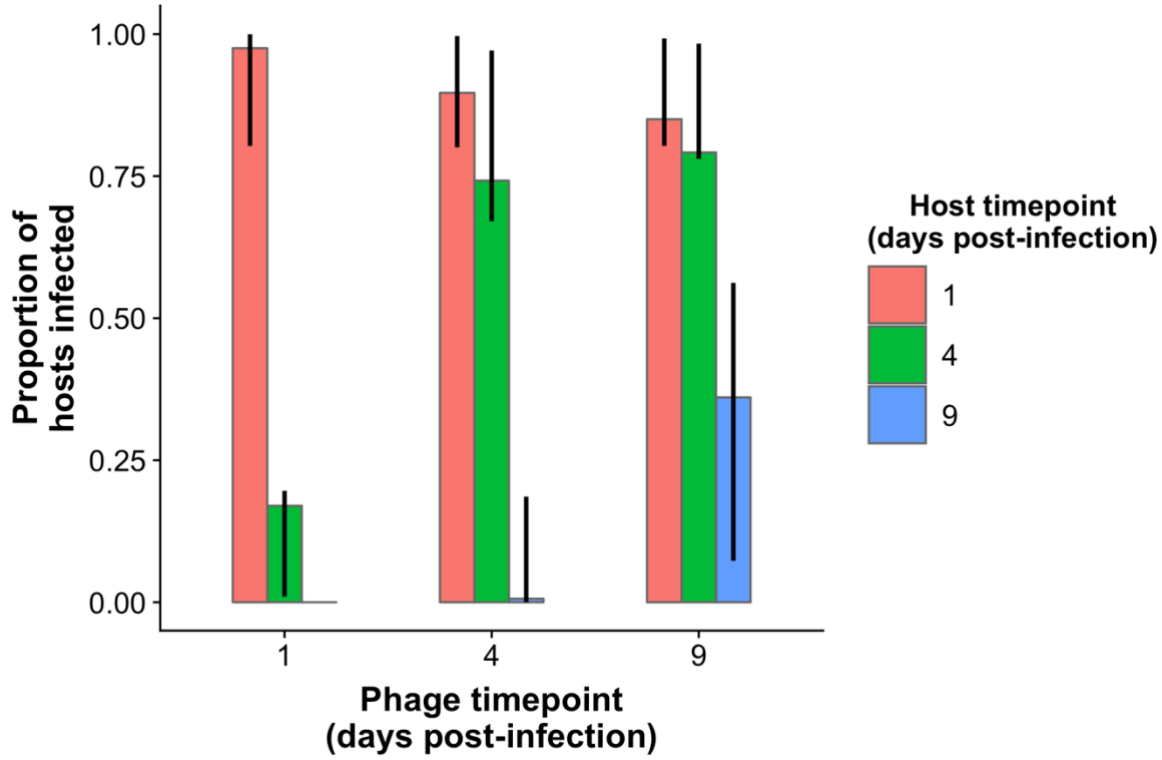


Figure A4.1 Proportion of hosts resistant to phage that were from the host's past, present or future from the time-shift experiment. Means and 95% CIs are shown (N=10048).

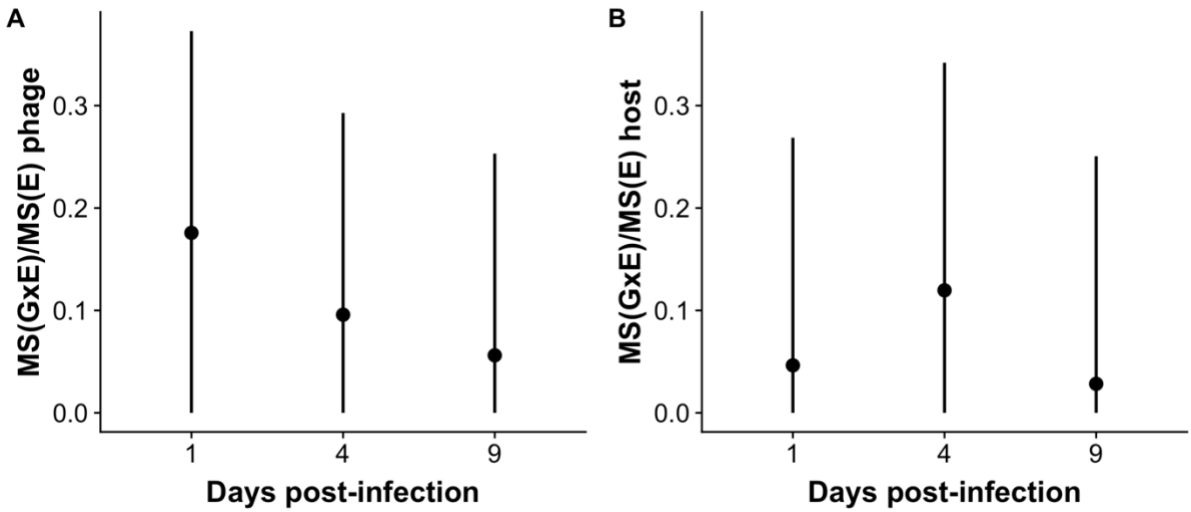


Figure A4.2 Relative importance of fluctuating selection dynamics (FSD) to arms race dynamics (ARD). Under FSD, pathogen genotypes should differ in their infectivity to hosts from contemporary and non-contemporary environments. Scores represent the ratio of the variance in infectivity due to host environment alone explained by variance in infectivity among genotypes, derived from GLMMs with environment as a fixed effect and phage genotype as a random effect. Residuals were square-root transformed to introduce a normal distribution in line with model assumptions. Means and 95% CIs are shown are shown (N=8).

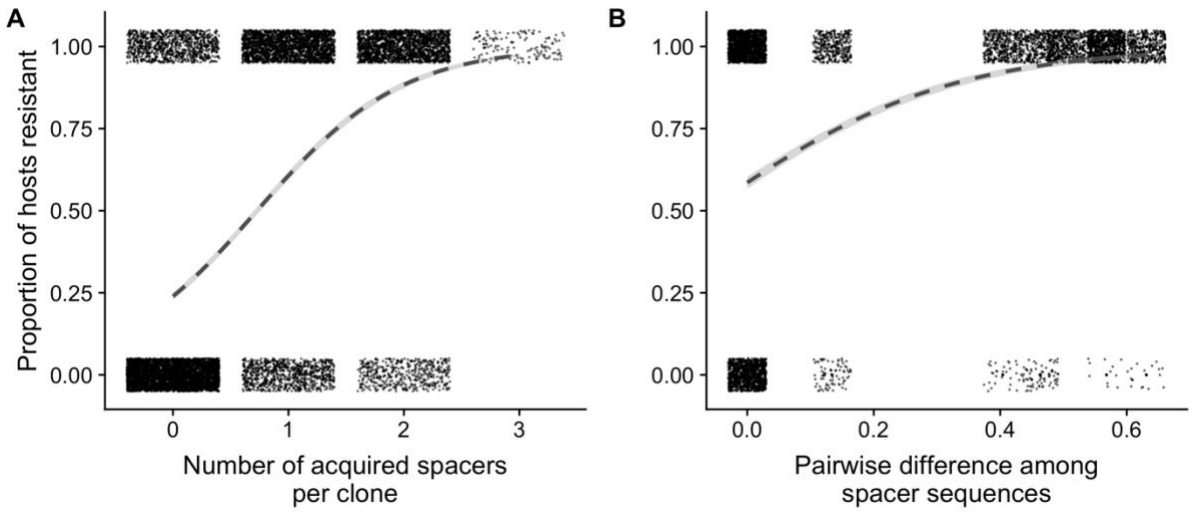


Figure A4.3 Relationship of host susceptibility to **A**) the number of spacers acquired per clone and **B**) sequence diversity in terms of the pairwise difference among spacer sequences from the phenotypic data. Points show raw data per clone, with random noise added to give a better indication of the number of occurrences of a given measurement. The dashed line is the smoothed logistic (binomial) regression slope fitted through the data. 95% confidence intervals are show in grey. N=10048

Supplementary Table Captions

Table A4.1

Infectivity matrices from the phenotypic assay (see *Experimental Methods*). The timepoint from which hosts and phage originated is shown. Each 12x12 block is a timeshift challenge within each replicate (1-8). In each block, columns 1-12 are individual bacterial clones and rows 1-12 are individual phage isolates. 1=plaque visible, measured infective; 0=no plaque visible, measured resistant. Phage clones selected for protospacer sequence analysis are indicated by black borders.

Table A4.2

Unique host CRISPR spacers detected by PCR analysis. *Replicate* and *Timepoint* (days post-infection) of the clone(s) which had a given spacer are shown. *Locus* is either CRISPR1 (CR1) or CRISPR3 (CR3). *Start* and *End* are the locations on the phage 2972 genome to which the sequences mapped, given in base pairs. *N* is the number of clones in a replicate X timepoint combination (max. 12) that had the spacer. *GeneID* is the NCBI number of the gene which the spacer mapped against, and refer directly to the. *Gene description* is the function (if known) of the region targeted by the spacer.

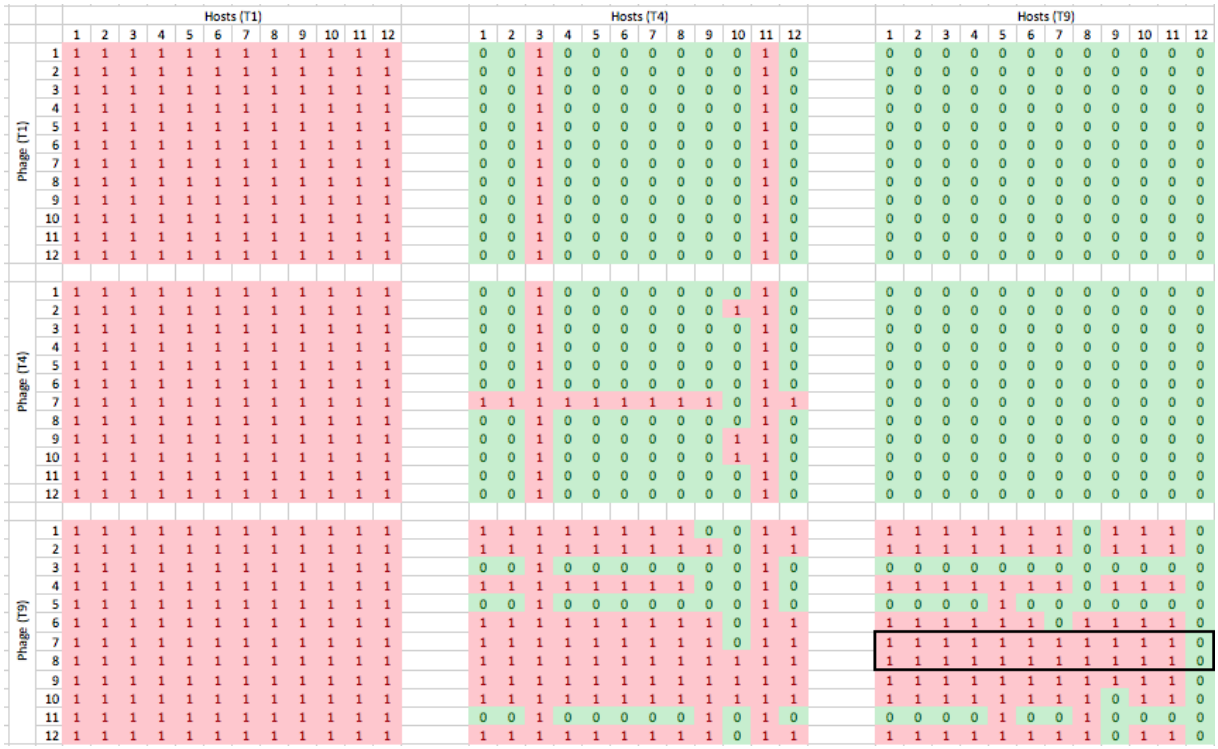
Table A4.3

Primer sequences used for protospacer analysis. *Annealing position* indicates where the primer binds to the phage 2972 genome.

Table A4.4

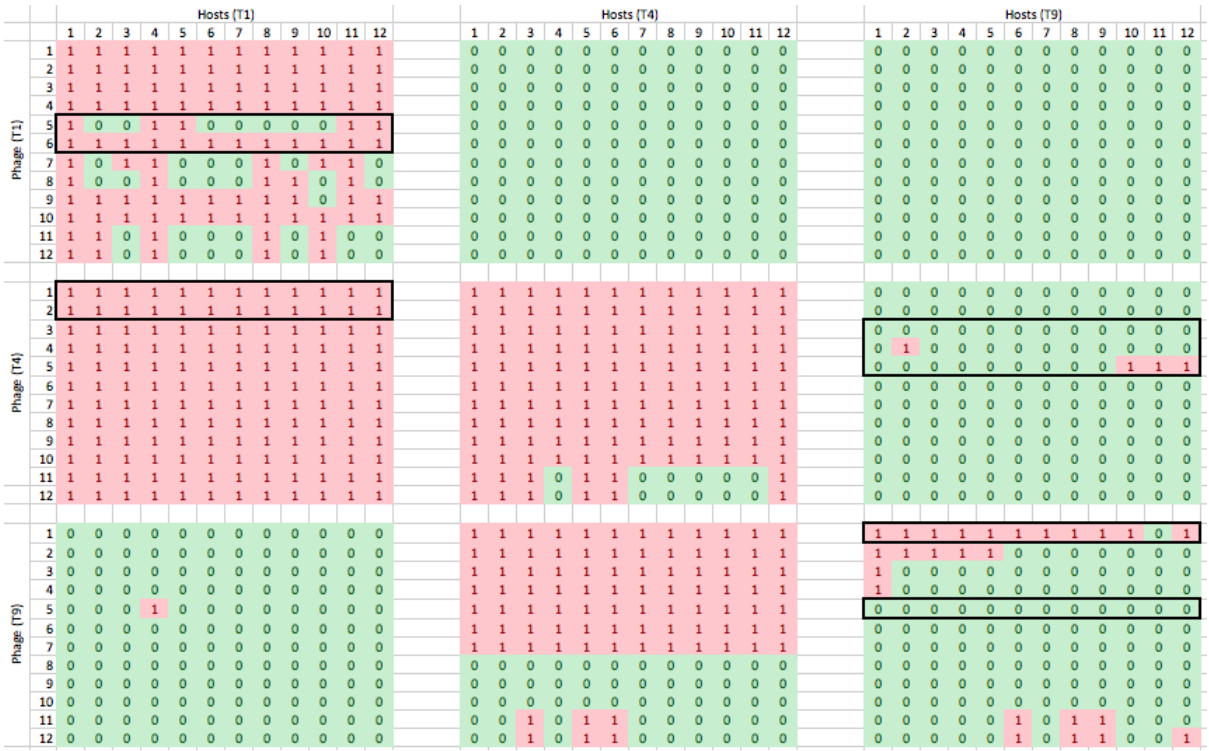
Location(s) of SNP(s) detected by PCR when mapped to the phage 2972 genome. *Replicate* and *Timepoint* (days post-infection) of the phage isolates are shown. *Phage ID* is the number of the phage isolate (rows 1-12 in the infectivity matrices [**Table A4.1**]). *SNP in protospacer* indicates if the SNP(s) was in the seed sequence or protospacer-adjacent motif of the target protospacer. Locations are given in base pairs.

Table A4.1



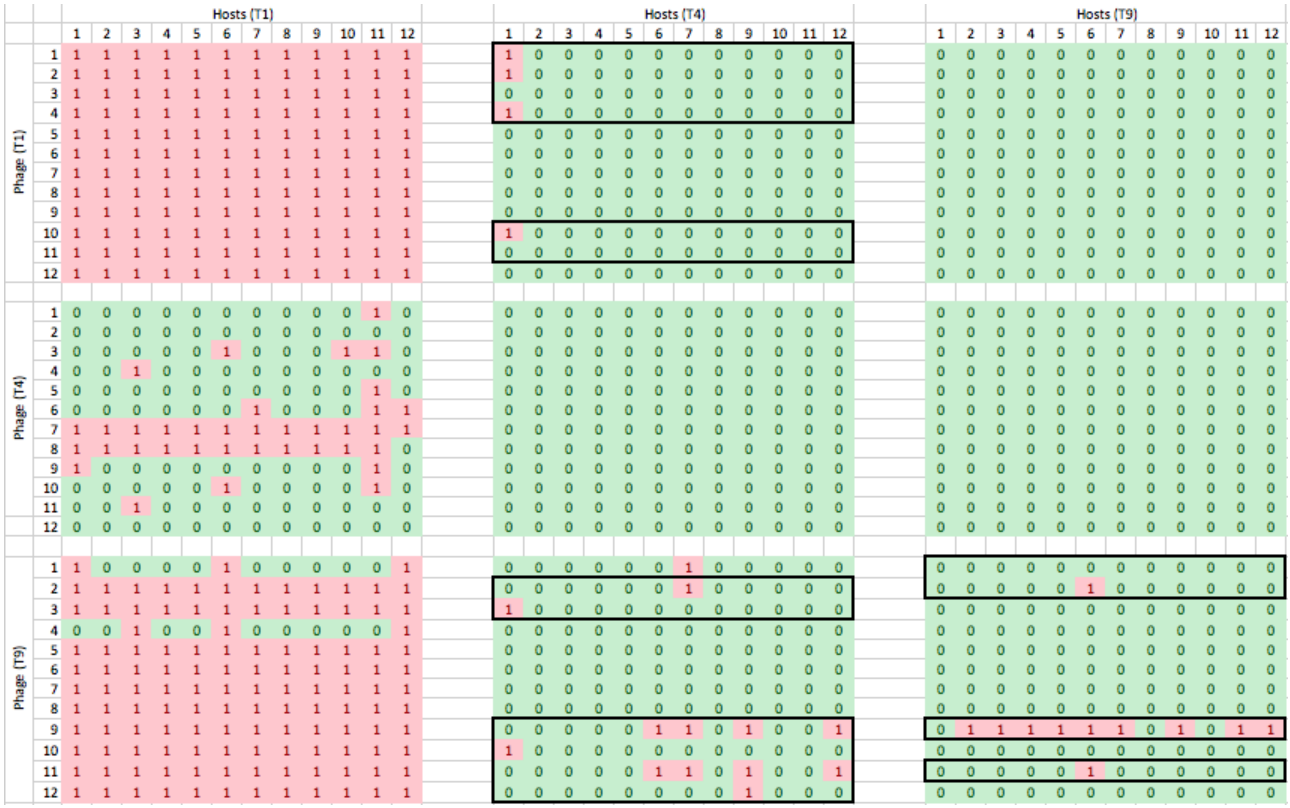
Phage	Hosts (T1)												Hosts (T4)												Hosts (T9)											
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Phage (T1) 1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0
3	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	0	1	1	1	0
5	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0
11	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	0

Replicate 1

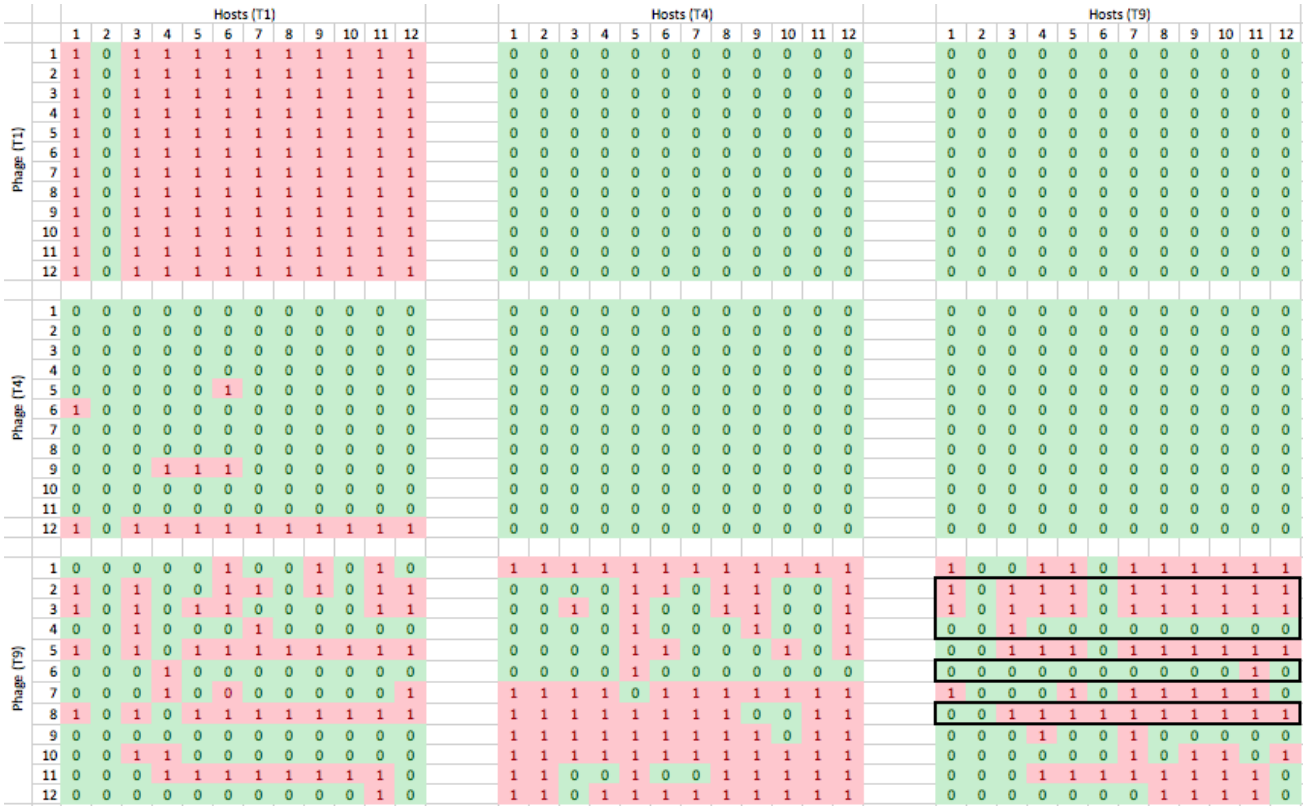


Phage	Hosts (T1)												Hosts (T4)												Hosts (T9)											
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Phage (T1) 1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	0	1	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	1	1	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1

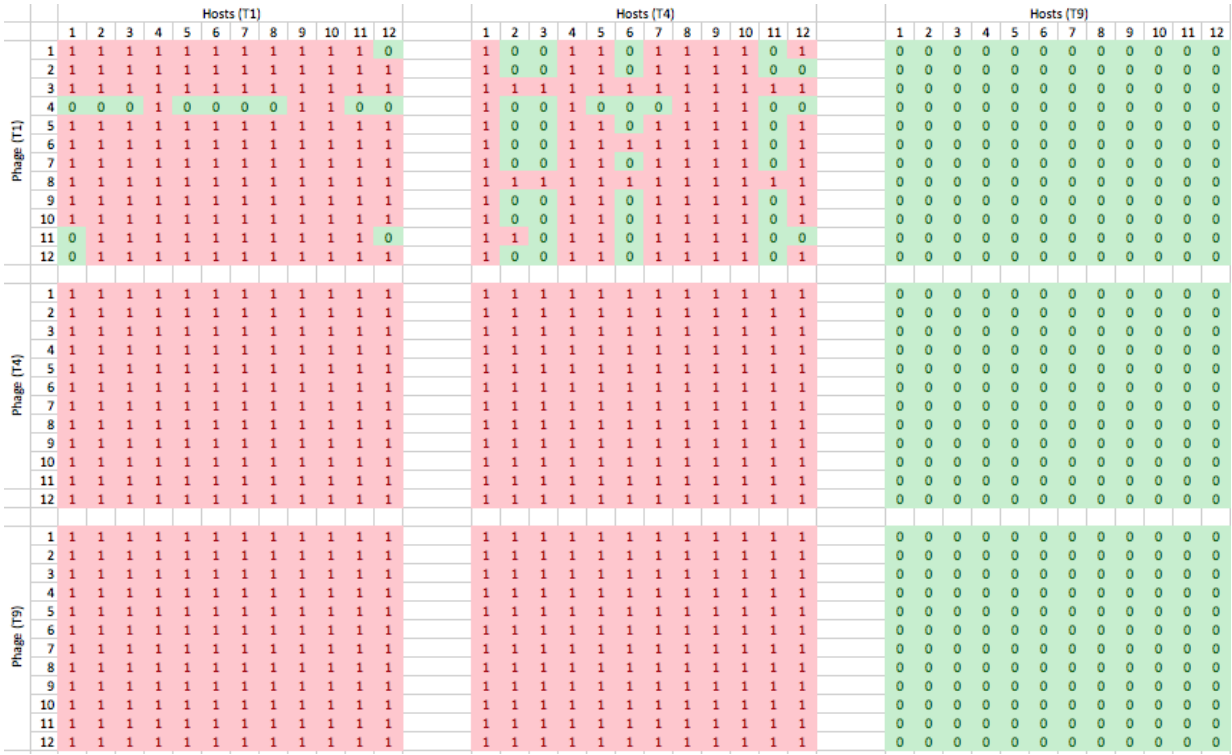
Replicate 2



Replicate 3



Replicate 4



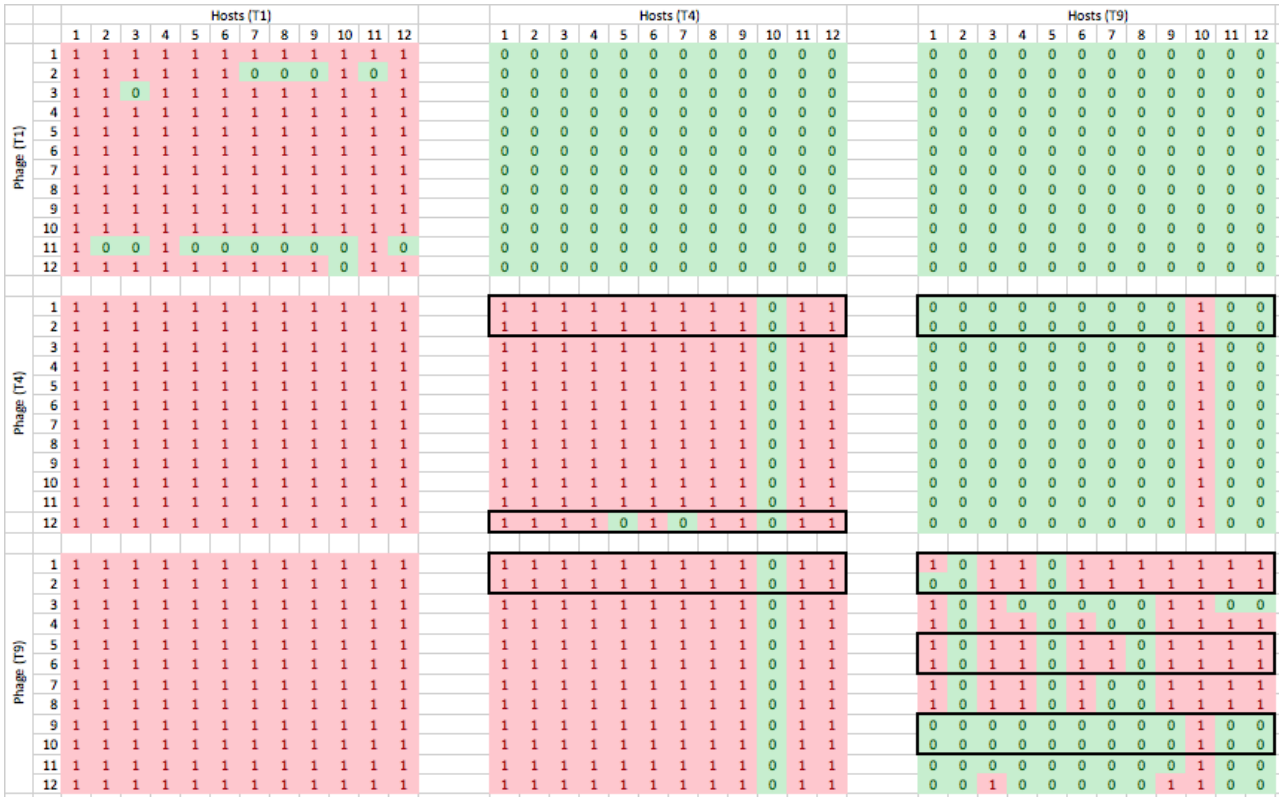
Phage	Hosts (T1)											
	1	2	3	4	5	6	7	8	9	10	11	12
Phage (T1) 1	1	1	1	1	1	1	1	1	1	1	1	0
Phage (T1) 2	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 3	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 4	0	0	0	1	0	0	0	0	1	1	0	0
Phage (T1) 5	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 6	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 7	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 8	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 9	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 10	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 11	0	1	1	1	1	1	1	1	1	1	1	0
Phage (T1) 12	0	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 1	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 2	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 3	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 4	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 5	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 6	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 7	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 8	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 9	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 10	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 11	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 12	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 1	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 2	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 3	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 4	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 5	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 6	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 7	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 8	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 9	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 10	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 11	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 12	1	1	1	1	1	1	1	1	1	1	1	1

Phage	Hosts (T4)											
	1	2	3	4	5	6	7	8	9	10	11	12
Phage (T1) 1	1	0	0	1	1	0	1	1	1	1	0	1
Phage (T1) 2	1	0	0	1	1	0	1	1	1	1	0	0
Phage (T1) 3	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 4	1	0	0	1	0	0	0	1	1	1	0	0
Phage (T1) 5	1	0	0	1	1	0	1	1	1	1	0	1
Phage (T1) 6	1	0	0	1	1	1	1	1	1	1	0	1
Phage (T1) 7	1	0	0	1	1	0	1	1	1	1	0	1
Phage (T1) 8	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T1) 9	1	0	0	1	1	0	1	1	1	1	0	1
Phage (T1) 10	1	0	0	1	1	0	1	1	1	1	0	1
Phage (T1) 11	1	1	0	1	1	0	1	1	1	1	0	0
Phage (T1) 12	1	0	0	1	1	0	1	1	1	1	0	1
Phage (T4) 1	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 2	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 3	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 4	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 5	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 6	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 7	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 8	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 9	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 10	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 11	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T4) 12	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 1	0	1	1	1	0	0	1	1	1	1	1	0
Phage (T9) 2	0	0	1	1	1	0	0	1	1	1	0	1
Phage (T9) 3	0	0	0	1	0	0	1	0	0	0	0	0
Phage (T9) 4	0	1	0	1	1	0	0	0	0	0	0	0
Phage (T9) 5	0	1	1	1	0	0	1	0	0	0	0	1
Phage (T9) 6	0	1	1	1	1	0	1	0	0	0	0	1
Phage (T9) 7	0	1	0	1	1	0	0	0	0	0	0	0
Phage (T9) 8	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 9	1	1	1	1	1	1	1	1	1	1	1	1
Phage (T9) 10	0	1	1	1	1	1	1	0	0	0	1	0
Phage (T9) 11	0	0	0	0	0	1	1	0	0	0	0	0
Phage (T9) 12	0	0	1	0	0	0	1	0	0	0	0	0

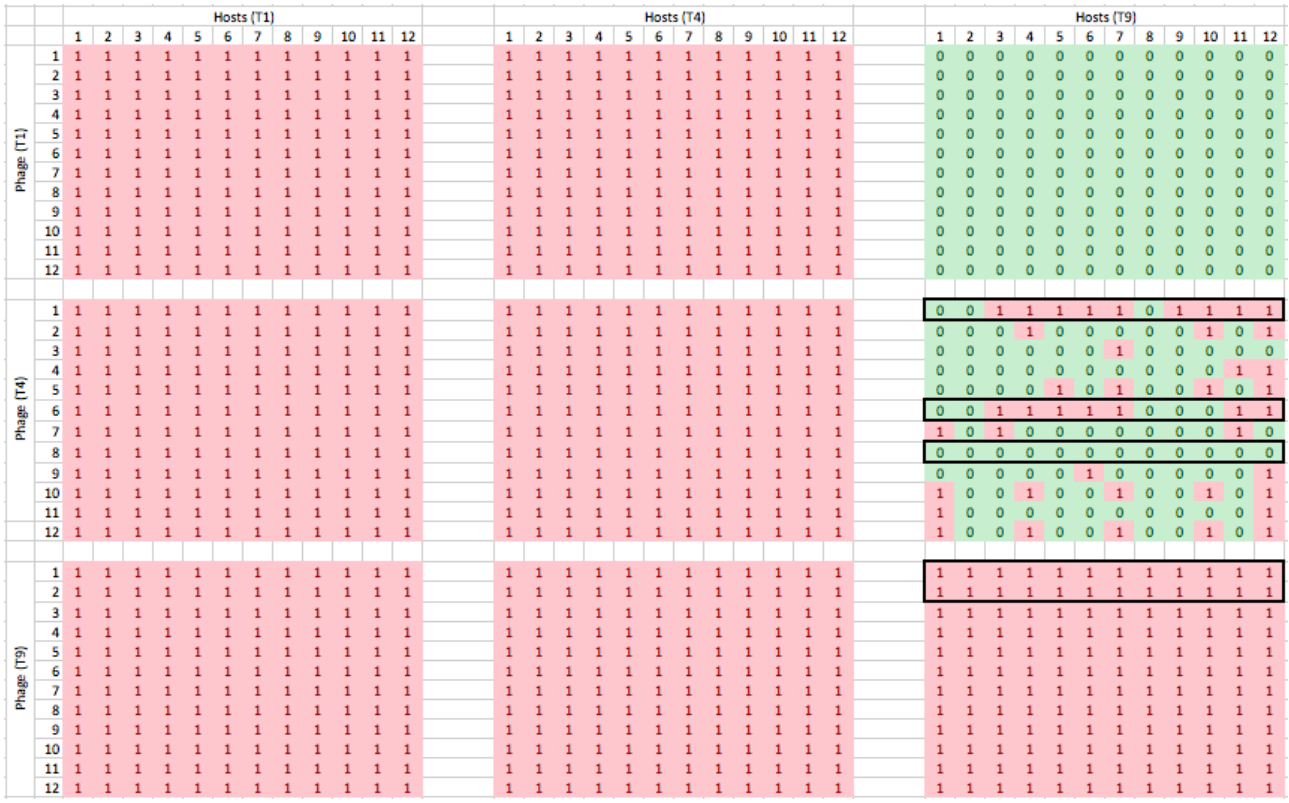
Phage	Hosts (T9)											
	1	2	3	4	5	6	7	8	9	10	11	12
Phage (T1) 1	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 2	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 3	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 4	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 5	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 6	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 7	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 8	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 9	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 10	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 11	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T1) 12	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 1	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 2	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 3	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 4	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 5	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 6	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 7	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 8	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 9	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 10	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 11	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T4) 12	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 1	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 2	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 3	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 4	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 5	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 6	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 7	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 8	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 9	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 10	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 11	0	0	0	0	0	0	0	0	0	0	0	0
Phage (T9) 12	0	0	0	0	0	0	0	0	0	0	0	0

Replicate 5

Replicate 6



Replicate 7



Replicate 8

Replicate	Timepoint	Locus	Start	End	Sequence	N	Gene description	GeneID
2	9	CR1	647	676	ATATCGTCCAGACTATCGCAGAATACTGAT	11	Hypothetical protein	5176920
3	4	CR1	871	900	GTTTTAAGTGGTATTATTATATTATCGAAG	1	Intergenic region	
3	9	CR1	871	900	GTTTTAAGTGGTATTATTATATTATCGAAG	10	Intergenic region	
3	4	CR1	1020	1049	CATAGCTTACAATGCGGCTCTTAAAGCTGG	4	Terminase small subunit	5176919
5	9	CR1	2214	2185	GGTCATCACCATTAACTAATCGAATAAGAT	1	Intergenic region	
3	9	CR1	2821	2850	TTTTTCGATGAGATATATCAGTACCGATGG	1	Terminase large subunit	5176917
6	4	CR1	10717	10746	ACGAAGGCTTGGAAAACATTTGATGGTAAC	12	Tail protein	5176903
6	9	CR1	10717	10746	ACGAAGGCTTGGAAAACATTTGATGGTAAC	12	Tail protein	5176903
2	9	CR1	16002	16031	TATGGTGTCGAAACTTTTAAACGTTACAAT	1	Tail protein	5176902
3	4	CR1	20696	20725	TGGAATAAGGTATCGTGCGTTTTGACAAGC	1	Antireceptor	5176901
6	9	CR1	22424	22395	CACATGATCTACAACTAGGTCAAGATTGCT	2	Structural protein	5176900
5	9	CR1	23363	23334	ATAGTTAACGCCTTTACACCGATCGAGGAA	5	Structural protein	5176900
4	4	CR1	23571	23600	TTGTTAAAAGAAGCACTAGAGGTGATTTAC	12	Hypothetical protein	5176899
4	9	CR1	23571	23600	TTGTTAAAAGAAGCACTAGAGGTGATTTAC	12	Hypothetical protein	5176899
6	9	CR1	25561	25531	AAACAAAAAATCTTTGAAGTTTATGACATAC	1	Hypothetical protein	5176893
5	9	CR1	26669	26639	GCATACCACAGGTATGACCAAAAACAAGAAA	2	Hypothetical protein	5176889
2	9	CR1	27003	27032	AGATTTATAACATGGAAATTGACGATGAAA	11	Hypothetical protein	5176888
6	9	CR1	27003	27032	TTTCATCGTCAATTTCCATGTTATAAATCT	2	Hypothetical protein	5176888
3	9	CR1	27358	27388	AACACTCAAAGAGTTACTTAAATCTGGAAAG	9	Hypothetical protein	5176888
6	9	CR1	27410	27380	CTGGAAAGCATATTGAGGGAGCTACTCTTG	2	Hypothetical protein	5176888
3	4	CR1	28744	28773	AGCTATAGTATATACACATAGCGTAGAAGC	3	Helicase	5176886
6	9	CR1	29124	29095	CAAAATAATTGTGGAAAATCACTGGTAAGT	1	Helicase	5176886
6	9	CR1	29407	29378	AACAACCATTATTTGGGTTGGCCCGAATAT	1	Helicase	5176886
3	9	CR1	29618	29647	TTTTCCGTCTTCTTTTTTAGCAAAGATACG	1	Hypothetical protein	5176885
4	9	CR1	29618	29647	TTTTCCGTCTTCTTTTTTAGCAAAGATACG	11	Hypothetical protein	5176885
5	9	CR1	29618	29647	TTTTCCGTCTTCTTTTTTAGCAAAGATACG	5	Hypothetical protein	5176885

2	1	CR1	29894	29923	TTATGGAGATGGTTGATTACGCAATCAACT	12	Replication protein	5176884
3	9	CR1	30017	29988	GCGACTGTTTGGTGGTTACTGACTTTTGCT	8	Replication protein	5176884
1	9	CR1	31582	31611	CTCAGTCGTTACTGGTGAACCAGTTTCAAT	12	Primase	5176883
3	9	CR1	32557	32584	GTATCAAACCAACGTCCATCAGCCATTC	1	Hypothetical protein	5176882
7	9	CR3	32554	32584	TCAGAATGGCTGATGGACGTTGGTTTGATAC	11	Hypothetical protein	5176882
7	4	CR3	32554	32584	TCAGAATGGCTGATGGACGTTGGTTTGATAC	12	Hypothetical protein	5176882
8	9	CR1	32707	32736	TGAAAAAACGAGGAGCACTCGTAGGAGTGG	12	Hypothetical protein	5176882
4	9	CR1	33769	33799	CTGCGTGGAACTGTCAGAACATAGTAGACTG	1	Hypothetical protein	5176878
6	9	CR1	33877	33848	TGGTAACTGAAAGGTCAGTGGAACGGCACG	1	Hypothetical protein	5176878
6	9	CR1	33998	33968	GTACAGAATTATTGAGGAGTTTATTGAACCT	5	Hypothetical protein	5176878
3	4	CR1	34587	34616	AGCCTAGATAGCGAAGTTGATCGTATCTAT	2	Hypothetical protein	5176877

Table A4.2

Table A4.3

Primer name	Primer sequence	Annealing position
1F	ACCAGTTGGAAGGAAAAGCTCT	29733
2F	AGACGGCTCATTTGTGGGTT	911
3F	CAACAGCAGCAAACACTGGG	20496
4F	GACAACGGAAACATCGCACC	28574
5F	GCACACCCTAACTGCGTCAT	34377
6F	TGGACCCCTAGCGGAAGTTA	23291
7F	ATGGCATGTCTAGCGCTCTC	10532
8F	GTGGGCACTGCTAAGAGTGT	32412
9F	TCAATCGGCTTTGAACGCAC	726
10F	CTGAACGTTTCGGTCTTGCC	31485
11F	GAATATCCACGCTGGCGAGA	474
12F	TTGGGGTCGTCCTCACATTG	15809
13F	AATTGAAGCACATCGGGGGA	26835
14F	AGCAAGGAAACTGACTGGCA	2625
15F	GCAGCGCTTGCGATTAGTAT	27206
16F	GTCCGATGTGTGGTCACGAA	29208
17F	TCAACCATTGGGCAGACGAA	33405
26F	AAGAGCGGTGTCCTCGAAAG	32230

Table A4.4

Timepoint	Replicate	Phage ID	SNP in protospacer	SNP location 1	SNP location 2	SNP location 3
1	2	5	1	29896		
1	2	6	1	29899		
4	2	1	1	29899		
4	2	2	0			
4	2	3	0			
4	2	4	0			
4	2	5	0			
4	3	1	0			
4	3	2	1	1039		
4	3	3	0			
4	3	4	1	1039		
4	3	10	1	1027		
4	3	11	0			
4	6	3	1	10724		
4	6	4	1	10724		
4	6	5	1	10724		
4	6	6	1	10724		
4	6	9	1	10724		
4	7	1	1	32580		
4	7	2	1	32583		
4	7	12	1	32580		
4	8	1	0			
4	8	6	0			
4	8	8	0			
9	1	7	1	31615		
9	1	8	1	31615		
9	2	1	1	682		
9	2	5	1	682		
9	3	2	1	893	1014	
9	3	3	0			
9	3	9	1	893	1023	29983
9	3	10	0			
9	3	11	1	893	1024	1027
9	3	12	1	893	1023	
9	4	2	0			
9	4	3	1	29614	29626	
9	4	4	1	29614		
9	4	6	1	29614		

9	4	8	1	29614	29623	
9	6	1	1	10724		
9	6	8	1	10724		
9	6	9	1	10724		
9	7	1	1	32587		
9	7	2	1	32587		
9	7	5	1	32587		
9	7	6	1	32583		
9	7	9	1	32587		
9	7	10	0			
9	8	1	1	32747		
9	8	2	1	32747		
Ancestral	NA	1	0			
Ancestral	NA	2	0			
Ancestral	NA	3	0			
Ancestral	NA	4	0			
Ancestral	NA	5	0			
Ancestral	NA	6	0			
Ancestral	NA	7	0			
Ancestral	NA	8	0			
Ancestral	NA	9	0			
Ancestral	NA	10	0			
Ancestral	NA	11	0			
Ancestral	NA	12	0			
Ancestral	NA	13	0			
Ancestral	NA	14	0			
Ancestral	NA	15	0			
Ancestral	NA	16	0			
Ancestral	NA	17	0			
Ancestral	NA	18	0			